BNL-96154-2011-JA

OBSERVING AND MODELING EARTH'S ENERGY FLOWS

Stevens, B. and Schwartz, S. E.

Submitted to
Surveys in Geophys.

July 2011

Atmospheric Sciences Division/Environmental Sciences Dept.

Brookhaven National Laboratory

U.S. Department of Energy
Office of Science

Surveys in Geophysics manuscript No.
(will be inserted by the editor)

Observing and Modeling Earth's Energy Flows

Bjorn Stevens · **Stephen E. Schwartz**

Received: date / Accepted: date

Abstract This paper introduces and summarizes, from the authors' perspective, some of the main challenges related to observing and modeling energy flows in Earth's climate system as discussed at a recent ISSI workshop entitled "Observing and Modeling Earth's Energy Flows". We review the state of understanding of Earth's energy balance and its susceptibility to perturbations, with particular emphasis on the roles of clouds and aerosols. Planetary albedo is rather narrowly constrained to 0.29 ± 0.01 by the balance of other components of the energy budget, including measurements of ocean heat uptake, but the surface energy budget cannot be closed because of large ($\simeq 15\ \mathrm{W\,m^{-2}}$) uncertainties in several terms. No evidence for secular trends in Earth's energy budget is found in ten years of CERES data, but the precision of the measurements is insufficient to meaningfully constrain forcings or feedbacks in the climate system. Efforts to conceptualize the system to isolate the role of clouds are reviewed, and differences between cloud radiative effects, forcings, and feedbacks are illustrated. Aerosol forcing remains the wildcard in forcing over the industrial era. Large regional trends in aerosol loading over the past decade exhibit no obvious association with regional trends in clear- or cloudy-sky radiative fluxes. Attempts to constrain climate sensitivity using estimates of radiative forcing, together with observed changes in globally averaged surface temperatures, are susceptible to large uncertainties in model estimates of aerosol forcing. The resulting uncertainties in sensitivity are likely to be as large as, or larger than, uncertainties in estimates of climate sensitivity from climate models. The utility of the climate sensitivity concept is reviewed; it is argued that, because of ambiguity in definitions and uncertainty associated with forcing induced by changes in greenhouse gases, the response of the climate system to an idealized perturbation in solar radiative forcing may be more useful in characterizing and comparing model sensitivities than the response to doubled CO_2, which has traditionally served as the benchmark in such studies.

Keywords Climate change · Cloud radiative effects · Aerosol · Energy budget · Climate sensitivity · Radiative forcing

1 Introduction

Modern climate science, wherein descriptive pictures of the climate system began to be complemented by quantitative theory, is only about a hundred years old. In the late 19th Century understanding of radiative transfer, particularly at the infrared wavelengths associated with terrestrial radiation, was developing rapidly, and it became possible to formulate quantitative descriptions of the relationship between the flux of energy through the Earth system, and quantities like the average surface temperature. Exemplary in this respect are Arrhenius' 1896 calculations suggesting that changes in carbon dioxide would induce changes in the surface temperature. Arrhenius' study was essentially an exercise in radiative transfer in which he quantified the flow of solar and terrestrial radiation through the earth system and the roles of the various processes that influenced these transports. Greenhouse gases such as carbon dioxide and water vapor played a key role in Arrhenius' calculations, and other factors influencing solar radiation such as clouds and surface properties were also accounted for. Although prescient in many respects, for instance with respect to the role of the

B. Stevens
Max-Planck-Institüt for Meteorology, 20147 Hamburg, Germany
Tel.: +49 40 41173-422
Fax: +49 40 41173-430
E-mail: bjorn.stevens@zmaw.de

S.E. Schwartz
Atmospheric Sciences Division, Brookhaven National Laboratory, Upton NY 11973, USA.

carbon cycle and feedbacks associated with water vapor and surface albedo, Arrhenius did not touch on a number of issues that have come to dominate the discourse with respect to climate change. Among these are heat uptake by the ocean and changes in the ocean circulation; changes in patterns of precipitation; the role of aerosols, both in influencing clouds and in the energy budget as a whole; and also the possibility that changes in cloudiness may enhance or offset other changes in the climate system.

An increased emphasis on precisely those issues largely left out of the early studies has marked a fundamental shift in climate science over the last forty or so years. Through the latter part of the 1960s climate science was preoccupied with resolving controversies related to the radiative transfer that was the basis for the analysis of Arrhenius and his followers. Confusion about the nature and importance of the details of the spectroscopy of CO_2 and H_2O in the thermal infrared was resolved only by the calculations of Manabe and Wetherald (1967), which showed that earlier controversies lost relevance when the vertical structure of the atmosphere is properly accounted for (cf., Pierrehumbert, 2011). Their research propelled the field into the current era, wherein qualitatively new questions, in particular the role of clouds and aerosol particles, demanded a more detailed understanding and accounting of the energy flows through the climate system.

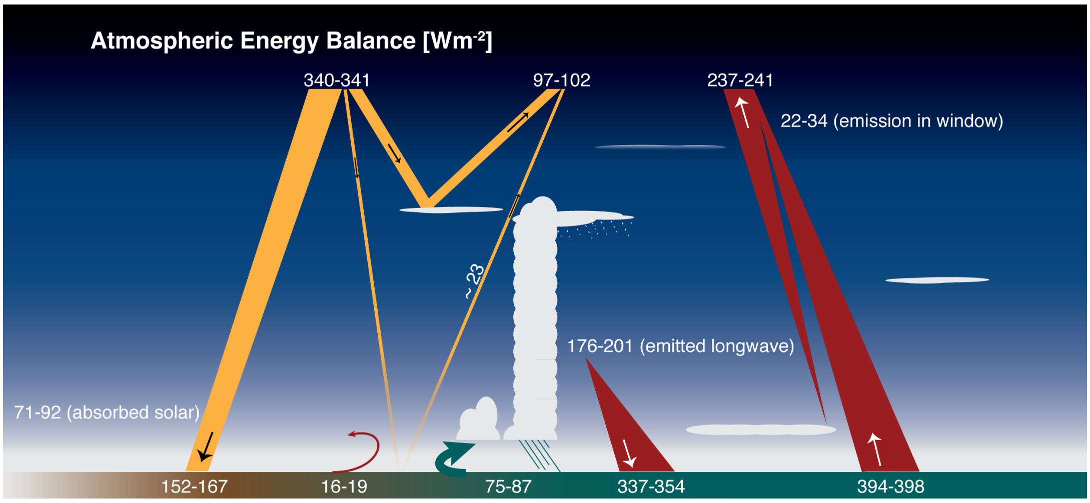

Fig. 1 Terms in Earth's global and annual mean top-of-the-atmosphere (TOA) and surface energy budget. TOA values are taken from ten years of CERES-SYN (synoptic) data, with the uncertainty range being set to plus or minus approximately twice the difference between the default estimates by CERES (the SYN product) and estimates adjusted to be consistent with an estimate of ocean heat uptake (the Energy Balanced and Filled, EBAF product). At the surface, radiative fluxes are taken from the last twenty years of the SRB (Surface Radiation Budget Stackhouse et al., 2011) data, with the uncertainty range adjusted to accommodate estimates based on ISCCP and (in the case of upward surface longwave fluxes) the reanalysis data. Turbulent surface flux data are taken from the ERA-Interim (European Center for Medium Range Weather Forecasts Reanalysis Interim product) data, with uncertainty estimates from precipitation incorporating measurements of that quantity based on precipitation climatologies as summarized by Trenberth et al. (2009).

What hasn't changed in the 115 years that have passed since Arrhenius' landmark study is the appreciation of the role of Earth's energy budget for the climate and more explicitly the recognition that comparatively small changes in flows of energy through the system can have a large impact. To lend some perspective we note that a doubling of CO_2, which is the paradigmatic example of a forcing of climate change, gives rise to a radiative perturbation of 3-4 $\mathrm{W\,m^{-2}}$ which is about 1% of the average flux of solar radiation incident on the Earth system;[1] the total radiative forcing attributed to long-lived greenhouse gases introduced through human activities over the industrial era thus far is about 3 $\mathrm{W\,m^{-2}}$. By way of comparison, in the tropics the diurnal variation of incident solar radiation is more than 1000 $\mathrm{W\,m^{-2}}$. The presence of a high cloud can change the outgoing terrestrial radiation by 100-200 $\mathrm{W\,m^{-2}}$, comparable to seasonal changes in radiative fluxes, albeit more short lived. From the perspective of climate system response, the central value of current estimates of the increase in global mean surface temperature that would result from a doubling of CO_2, 3 K, is about 1 % of the global mean surface temperature, 288 K, and again, much less than geographical and temporal variability. Given that internal changes to the system in response to a doubling of CO_2 can be considered significant if they perturb the system by such a small amount relative to both mean values and spatio-temporal fluctuations, it is perhaps not surprising that it has proven difficult to construct models, or collect observations, that can substantially improve quantitative understanding

[1] Unless otherwise noted the radiative forcing estimates we cite are for the radiative perturbation that arises at the TOA after the stratosphere has adjusted to the presence of compositional change, but with the tropospheric temperatures held constant.

relative to what has been achieved by studies that did not elaborate the role of diverse feedback mechanisms. In this context it becomes legitimate to ask whether, for example, models that grossly misrepresent the diurnal cycle of clouds can be trusted to adequately represent the response of cloudiness to surface temperature changes expected to accompany a doubling of CO_2. Likewise, from a measurement perspective, it seems legitimate to ask whether it is possible even to measure the energy balance in the climate system with the precision required to track such changes? And if not, what are the prospects for doing so? Questions such as these, which are central to current climate research, served as the basis for discussions at the recent ISSI workshop on Earth's Energy Flows.

In the following we reflect on some of the issues that arose through the course of the discussions, emphasizing common threads that emerged in the fabric of disparate discussions. Our presentation is organized around three basic issues: (i) what is the status of present understanding of the Earth's energy budget; (ii) how does the composition of the atmosphere, particularly clouds and aerosols, influence this budget; and (iii) to what extent do models provide an adequate description of processes regulating the flow of energy through the climate system. These issues, particularly those aspects central to advancing understanding of the climate system, are discussed in turn below. Although the discussion presented herein benefitted greatly from presentations and discussions at the Workshop, responsibility for the material presented rests with the current authors and should not be taken as representing the views of participants in the workshop or as a workshop consensus.

2 Present understanding of Earth's energy budget

2.1 Global energy balance

The major pathways through which energy flows through the climate system are commonly summarized in figures such as Fig. 1, which have a long history in climate science. The first such figure is attributed to Dines (1917), although the measurements made and published by Abbot and Fowle (1908) tell a similar story, just without the diagram. Similar figures by London (1957) culminated the state of understanding of the Earth's energy budget at the dawn of the satellite era, and a figure presented by Ramanathan (1987) summarized understanding that had been gained in the early years of the Earth Radiation Budget Experiment (ERBE). In recent years the summary figures by Kiehl and Trenberth (1997) and Trenberth et al. (2009) have become standard references. Fig. 1 differs from that of Trenberth et al. (2009) in that it endeavors to present a consistent picture of not only current understanding of the energy flows through the climate system, but uncertainty in individual terms as well.

At the top of the atmosphere (TOA) the incoming solar irradiance is now confidently thought to be at the lower end of the range of previous estimates, and with a greatly reduced uncertainty (Kopp and Lean, 2011). The value and the reduced uncertainty result from identification of artifacts in the older measurements that lends enhanced credence to the newer, and lower estimates.

Although there is a rather larger inherent uncertainty in measurements of the reflected shortwave flux, this uncertainty can be reduced by demanding consistency with the other components of TOA energy flux. Here two additional terms come into play, the emitted longwave radiation, and the net flux imbalance at the TOA. The former is assigned a value of 239 ± 2 $\mathrm{W\,m^{-2}}$ and thus is relatively well measured, although some ambiguity arises in the daytime due to the overlap between the solar and terrestrial contributions in the near infrared; the latter can be inferred from measurements of the changing ocean heat content, and to a lesser extent melting of the cryosphere and warming of the land surface. As discussed by Lyman [this issue] the rate of heating of the top 700 m of the world ocean over the last sixteen years, as inferred from temperature measurements and expressed per the area of the entire planet, is 0.64 ± 0.11 $\mathrm{W\,m^{-2}}$ (90% confidence interval). Sparser measurements extending to ocean depths of 3 km reported by Levitus et al. (2005) suggest that the upper ocean takes up about three-quarters of the ocean heating. Levitus et al. (2005) also estimate that the contributions of other heat sinks, including the atmosphere, the land, and the melting of ice, can account for an additional 0.04 $\mathrm{W\,m^{-2}}$ Based on these measurements the flux imbalance at the TOA is estimated to be 0.9 ± 0.3 $\mathrm{W\,m^{-2}}$ The uncertainty is based on the 90% confidence interval given by Lyman and the assumption that the relative uncertainty in the deep ocean heat-uptake estimates and in the estimates of heating by other components of the Earth system are about 50%.[2] For reference, in constructing the energy-balanced version of the CERES data Loeb et al. (2009) estimated the surface heat uptake to be 0.85 $\mathrm{W\,m^{-2}}$ similar to the 0.9 $\mathrm{W\,m^{-2}}$ estimated here and employed by Trenberth et al. (2009) based on a somewhat different line of reasoning. Combining this estimate of the imbalance in radiative fluxes at the TOA with estimates of the emitted longwave radiation and incoming shortwave irradiance leads to an estimate of the reflected TOA shortwave radiative flux between 97 and 102 $\mathrm{W\,m^{-2}}$, which puts the corresponding albedo, 0.29 ± 0.01, at the lower end of historical estimates (cf. Hunt et al., 1986; Ramanathan, 1987; Trenberth et al., 2009), and reduces the uncertainty relative to what can be deduced from an analysis of measurements of reflected shortwave irradiance alone.

[2] Other sources of energy — geothermal, combustion of fossil fuel and nuclear production — are yet an order of magnitude smaller (Pilewski, this issue).

At the surface, uncertainties are much larger and distributed more uniformly over several terms. Based on a survey of the existing literature (particularly Stackhouse et al., 2011), tabulated values in Trenberth et al. (2009), discussions at the workshop, and the requirement of consistent uncertainty estimates when the heat budget of the atmosphere or surface as a whole is accounted for, we estimate an uncertainty of about 15 $\mathrm{W\,m^{-2}}$ in the shortwave flux absorbed by the surface, the downwelling longwave flux absorbed by the surface, and precipitation/evaporation. As pointed out by Trenberth et al. (2009), if each of the terms in the surface energy budget is estimated individually, in isolation of the others, an imbalance in the net surface flux of as much as 20 $\mathrm{W\,m^{-2}}$ can arise; this is more than an order of magnitude greater than current measurement based estimates of the rate of heat uptake by the ocean and land. Trenberth et al. (2009) addressed this issue by assuming that the downwelling longwave radiation carried the greatest uncertainty and then estimated it as a residual from the other terms, which included an estimate of the heating rate of the ocean and land. This procedure lead to an estimate of 333 $\mathrm{W\,m^{-2}}$ for the downwelling longwave radiative flux at the surface, a value which is outside of the range of our estimates. This discrepancy points to the need for the uncertainty to be carried by more terms than just the downwelling longwave radiative flux. Surface moisture fluxes estimates based on reanalysis products for evaporation versus satellite climatologies of precipitation, both of which are rather indirectly measured, differ considerably, as do published estimates of solar radiation absorbed by the surface. Even the lowest estimates for net radiation at the surface, 91 $\mathrm{W\,m^{-2}}$ imply a global average moisture flux larger than 0.9 m $\mathrm{yr^{-1}}$, which is larger than previous estimates. Hence the magnitudes of the uncertainties in all of these terms are comparable. Although the particulate component of the atmospheric composition (clouds, aerosols) is increasingly being constrained by passive and active remote sensing (Stephens et al., 2008; Winker et al., 2007), thus extending the hope that the surface energy budget may eventually be constrained by satellite measurements, it seems unlikely that the uncertainty in this budget will be meaningfully reduced in the near future without substantial augmentation of the spatial coverage of surface radiation measurements.

Much of the uncertainty in estimates of the energy budget, both at the TOA and at the surface, can be attributed to variability in atmospheric composition, mainly water vapor, clouds and aerosols, but particularly clouds, in view of their large spatial and temporal variability and the large changes they induce in short- and longwave radiation. As the CERES instruments used to derive the top of the atmosphere radiation budget are mounted on two polar orbiting satellites, with mid morning and early afternoon equator crossing times, they poorly sample the diurnal cycle; further, as these are narrow-field-of-view instruments they have difficulty determining the full angular distribution of radiation that contributes to short- and longwave flux. Twice daily shortwave radiance measurements (from each of the two platforms) are first converted to irradiance using scene-dependent angular distribution functions which are then adjusted, empirically, based on information from geostationary satellites to account for diurnal variability. However this procedure, while attempting to solve the problem of diurnal variability that is not resolved by the highly calibrated polar-orbiting instruments introduces uncertainty and possible bias (Loeb et al., 2003). Moreover the wide variability of cloud types in Earth's climate system makes it difficult to aggregate these effects in a small number of idealized scenes, models for which are used to convert radiances into fluxes. These limitations on the TOA budget might be partially overcome by use of calibrated broad-band radiation budget measurements from geostationary satellites tasked with monitoring a limited geographical region through the GERB (Geostationary Earth Radiation Budget) program (Harries et al., 2005). A potentially attractive alternative would be to deploy an array of sensors on the forthcoming Iridium fleet of 66 polar orbiting communications satellites. Still, at the end of the day, measurements are only beginning to approach the accuracy required to test understanding of the changing climate at the TOA, and remain woefully inadequate at the surface.

2.2 Radiative forcing, response and climate sensitivity

A consideration of the energy budget also provides the framework for understanding climate change. This framework, which has developed over the last thirty years, involves the assumption that changes in the climate system can be well characterized by linearly relating changes in the globally averaged surface temperature to a radiative forcing, F. The constant of proportionality between the forcing and the response is called the equilibrium climate sensitivity, S_{eq}, which can be formally defined as the steady-state change in T_{s} (the globally averaged near-surface air temperature) that would result from a sustained change in a radiative flux component of the Earth energy budget at the TOA (forcing), normalized to that flux change, with unit: K $(\mathrm{W\,m^{-2}})^{-1}$. That is, the equilibrium sensitivity is the proportionality constant between the steady-state change in surface temperature and the applied forcing.

$$\Delta T_{\mathrm{s}} = S_{\mathrm{eq}} F \tag{1}$$

Because it codifies the sensitivity of such an essential feature of the climate state, the globally averaged surface temperature, Eq. (1) provides a powerful description of past and prospective future climate change. To the extent that other properties of the climate system scale with its value, S_{eq} assumes an even broader significance. For these reasons, determination of S_{eq} has evolved into a central focus of climate science (*e.g.,* Knutti and Hegerl, 2008). Uncertainty in the way in which clouds respond to, and thereby mediate or amplify, forced changes to the state of the climate system underlies most of the uncertainty in S_{eq}; this has come to be known as the cloud-feedback problem.

Central to the conceptual framework resulting in Eq. (1) are the assumptions: (i) that S_{eq} does not depend on nature or magnitude of the forcing; and (ii) that the forcing resulting from a given change in atmospheric composition or surface properties can be determined unambiguously. Both assumptions are problematic. Models have long suggested that, S_{eq} depends on both the temperature and the nature of the forcing (*e.g.,* Hansen et al., 1997; Colman and McAvaney, 2009). For instance Hansen et al. (1997) showed that if the TOA radiative forcing is equated with the initial radiative perturbation associated with a change in CO_2, the change in T_s will differ by as much as 25% as compared to what arises for an equal radiative perturbation arising from a change in the solar-irradiance. To minimize such dependencies they developed the concept of an effective, or adjusted, forcing. The adjusted radiative forcing attempts to account for a rapid, and perturbation dependent, adjustment of the climate system to a compositional change. The familiar example, and the one that motivated the original concept of an adjusted radiative forcing, is the adjustment of the stratosphere which proceeds differently depending on whether the initial radiative perturbation is caused by changes in well-mixed greenhouse gases as compared to changes in the solar irradiance, or stratospheric aerosols (*e.g.,* Hansen et al., 1997). Less well appreciated and more difficult to understand, is how changes in cloudiness also contribute to ambiguity in the definition of an adjusted forcing (Hansen et al., 1997; Gregory and Webb, 2008; Andrews et al., 2009). Thus, because of the interactions of clouds with forcing agents, the use of Eq. (1) as a description of the climate system, may result in a poor understanding and quantification not only of the response of the climate system, S_{eq}, but even of the forcing, F.

A possible means of sidestepping many of the issues discussed above is offered by estimating climate sensitivity from observations. Although S_{eq} cannot be directly measured from a change to a new equilibrium (steady state) climate, because of the impossibility both of maintaining a sustained forcing relative to some well defined prior atmospheric state and of waiting for the global mean temperature to achieve a new temporally constant value, a consideration of the energy flows of the climate system suggest possible empirical approaches to its determination. Here it proves useful to explore the behavior of the system as it approaches a new steady state. Assuming that the forcing, F were known, it could be related to the rate of change in the heat content of the climate system, $\dot{H}$ and the temperature change between that at $t = 0$, defined to correspond to a time when the system is in stationarity and the forcing is applied, and that at some later time, t,

$$\dot{H} = F - S_{\mathrm{eq}}^{-1}(T_{\mathrm{s}}(t) - T_{\mathrm{s}}(0)). \tag{2}$$

Initially upon imposition of the forcing, T_{s} has not yet responded to the forcing, and the heat content of the planet increases in response to the forcing (taken here as positive). For a constant forcing F, as the surface temperature increases in response to the forcing, the rate of increase of the planetary heat content decreases as the amount of heat radiated from the planet increases in proportion to the increased temperature. Ultimately a new steady-state is reached whereby $\dot{H} = 0$, and

$$T_{\mathrm{s}}(\infty) = T_{\mathrm{s}}(0) + FS_{\mathrm{eq}} \tag{3}$$

Consequently if both the forcing and the rate of change of the planetary heat content are known the equilibrium sensitivity might be determined (Gregory et al., 2002) as

$$S_{\mathrm{eq}} = \frac{T_{\mathrm{s}}(t) - T_{\mathrm{s}}(0)}{F - \dot{H}} \tag{4}$$

where the forcing and the change in global temperature are relative to a prior unperturbed state (at time, $t = 0$) and the rate of heating of the planet $\dot{H}$ can be inferred either from satellite measurements at the top of the atmosphere or from the change in heat content of the planet inferred from ocean calorimetry as discussed elsewhere in this special issue (Lyman). As noted by Gregory et al. (2002) the uncertainty in forcing over the period of instrumental temperature record, due mainly to uncertainty in aerosol forcing, results in large uncertainty in the denominator of (4), and hence in uncertainty in the inferred sensitivity too large to be useful for any practical application. While the importance of aerosols in this respect has become widely appreciated, the role of clouds, which respond differently to aerosol versus greenhouse gas perturbations and thus help determine how a compositional change is converted into a radiative forcing F, is not widely recognized.

An alternative approach (Forster and Gregory, 2006; Murphy et al., 2009) is to determine S_{eq} over shorter time periods, not starting with the unperturbed state, as

$$S_{\mathrm{eq}} = \frac{T'}{F' - \dot{H}'} \tag{5}$$

where primes denote changes in the quantities over a given time period. Again the utility of any such determination of sensitivity rests on the uncertainties in the several quantities on the right hand side of the equation. By focusing on periods over which the observational record is better constrained one could presumably limit such uncertainty, although so doing does not mitigate the difficulties the cloud response plays in converting a change in atmospheric composition (or land surface properties) to a radiative perturbation. Moreover, by focusing on shorter periods both the numerator and denominator of Eqs. (4)-(5) become small; but in this case, because the denominator is the difference of two rather large terms, the answer becomes proportionally more sensitive to errors in either F or $\dot{H}$.

In summary, an inability to characterize past changes in the aerosol limits present ability to quantify the past radiative forcing of the climate system, and a poor understanding of cloud processes limits understanding of the response of the climate system to changes in its surface temperature. These points are generally well appreciated. Less well appreciated is that a poor understanding of cloud processes limits ability to convert changes in atmospheric composition, or surface properties, into radiative forcings and thus undermines the utility and generality of the climate sensitivity framework as a whole.

2.3 Trends

Monitoring changes and trends in these energy flows in the climate system, and in particular in cloudy and cloud-free regions, presents additional opportunities to develop understanding. For examination of trends the precision of the measurements is more important than the accuracy, and the CERES instrumentation was specifically designed also with this goal, of estimating changes in the energy budget in mind (Wielicki et al., 1996). However, the ability to use the CERES measurements in this way was recently called into question by Trenberth and Fasullo (2010b), who suggested that the radiative imbalance at the top of the atmosphere measured by CERES has been increasing by as much as 1 $\mathrm{W\,m^{-2}}$ per decade. Because of the limited heat capacity of the atmosphere such an imbalance would imply a large change in ocean heat content and/or surface ice amount, neither of which is observed. Trenberth and Fasullo (2010b) did not, however, account for the uncertainty in their estimated trend and made use of preliminary CERES data for the period (2005-2010), over which time period the striking trend in radiative imbalance was noted. This situation motivated us to revisit this question using a recently released version (edition 2.5) of the CERES data, for which calibrated measurements are available through 2010.

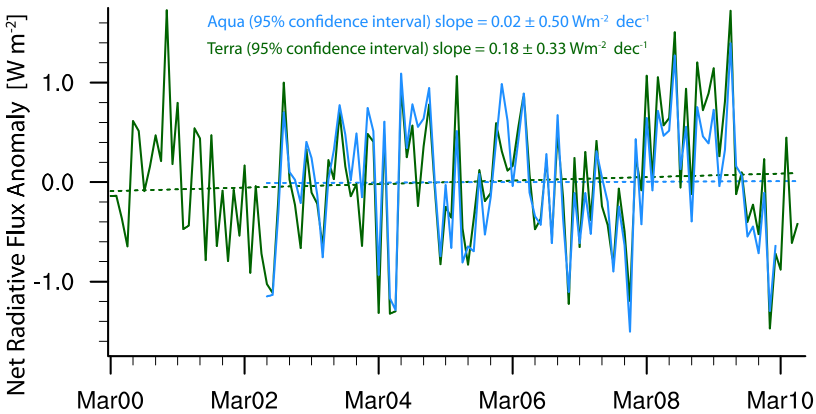

Fig. 2 Anomalies in the monthly and globally averaged top of the atmosphere net radiation from the CERES SSF1deg product. Anomalies are calculated by subtracting the CERES monthly climatology from the monthly fluxes. Regression line uncertainty is estimated assuming that monthly anomalies are independent. The standard deviation of the anomalies is about 0.65 $\mathrm{W\,m^{-2}}$.

This analysis is based on the SSF1deg (single scanner footprint at one degree) data set, which includes information only from the twice daily measurements of the CERES platform on the polar orbiting Terra (am) and Aqua (pm) satellites. Unlike the CERES SYN or EBAF products, the SSF1deg data set does not incorporate measurements from a shifting complement of geostationary satellites. For this reason it might be expected that this data set is more precise, as it more fully represents the inherent stability of the CERES measurements themselves. To explore this data set for trends the 1° gridded data from the SSF1deg product was averaged over the globe and over each month of measurements to create a monthly climatology. Monthly anomalies were constructed by subtracting this monthly climatology from the global and monthly averaged values. The analysis (Fig. 2) provides little evidence of a trend. Indeed, the data cannot confidently rule out the possibility of a negative trend, *i.e.,* in the opposite direction of the trend reported by Trenberth and Fasullo (2010b).

Ideally it is desired to measure the flow of energy through the Earth system accurately enough to allow for determination of the net planetary imbalance that is meaningful in the context of understanding the consequences of radiative forcings and precisely enough to allow for the identification of trends on the decadal time scale. So doing would aid understanding of how the system changes in response to various forcings, ranging from solar variability or volcanism, to the steadily increasing concentration of greenhouse gases. From the perspective of accuracy it should be noted that the CERES instrumentation indicates a net planetary energy imbalance at the TOA of 6.5 $\mathrm{W\,m^{-2}}$ (Loeb et al., 2009). Such an imbalance is certainly a measurement artifact and not a true imbalance. However, from the perspective of understanding the consequences of forcings on the planetary energy budget and for confident determination of climate sensitivity an accuracy that is more than an order of magnitude better is essential. From the perspective of trends the present analysis suggests that natural variability and limited sampling alone contribute to an uncertainty of 0.33 $\mathrm{W\,m^{-2}}$ per decade in the estimate of trend in the net radiative balance and hence that no meaningful conclusion can be drawn at present regarding the magnitude or sign of any trend in this imbalance.

3 Clouds

3.1 Radiative effects, feedbacks and forcings

A framework for considering cloud feedbacks in the climate system, first formulated by Schneider (1972), introduced the idea of cloud forcing (see also Cess, 1976; Ramanathan, 1987). Identifying the subset of short- and longwave radiative fluxes associated with cloud-free scenes permits defining ”cloud radiative forcing” as the difference between the all-sky or actual radiative flux and the contribution from scenes having skies determined to be cloud free. Here, as before, the term forcing is used to denote a change in radiative flux due to a change in atmospheric composition, namely clouds. However, because this terminology does not allow one to distinguish the radiative effect of the totality of clouds, from the radiative perturbation that would accompany a perturbation in cloudiness, it proves useful to distinguish between the cloud-radiative effect, CRE, as the radiative effect of the background state of cloudiness, and reserve the phrase “radiative forcing” for radiative perturbations driven by externally imposed changes.

The definition of the cloud radiative effect can be made precise as follows. Let Q denote the rate of absorption of solar energy and E denote the rate of emission of infrared energy, both at TOA. If an inward directed flux is defined to be positive, E at the TOA must be negative. With this sign convention and letting H denote the enthalpy of the Earth system, including components such as oceans, land and surface ice, then the rate of change of H with time[3], $\dot{H}$ is given by $\dot{H} = Q + E$. Under the assumption that three-dimensional radiative effects are negligible the two radiative flux terms may be conceptually distinguished into components pertaining to contributions from cloudy and cloud-free regions of the planet:

$$Q = Q_*\left[1 - \alpha_0(1 - A_\mathrm{c}) - \alpha_\mathrm{c} A_\mathrm{c}\right] \tag{6}$$
$$E = -E_0(1 - A_\mathrm{c}) + E_\mathrm{c} A_\mathrm{c}. \tag{7}$$

Here A_c denotes the fraction of the area of the planet that is, on average, cloudy; Q_* denotes the surface-averaged solar irradiance ($Q_* \approx 340\ \mathrm{W\,m^{-2}}$, *e.g.,* Fig. 1); α_0, and α_c denote the *effective* TOA albedo of the cloud-free and cloudy scenes, respectively; and E_0, and E_c denote the emitted longwave irradiance obtained by compositing over the cloud-free and cloudy scenes, respectively. The albedos are designated as “effective” quantities because Eq (6) depends non-linearly on cloudiness and insolation, and thus the effective albedos must account for the co-variability between these two quantities. So for instance, α_c is not the average cloud albedo, but rather the cloud albedo that the average cloudiness requires so that the planetary albedo, $\alpha = \alpha_0(1 - A_\mathrm{c}) - \alpha_\mathrm{c} A_\mathrm{c}$ matches that observed, likewise for the α_0, the effective albedo of the cloud free scenes. In general each of the terms on the right hand side of Eq. (6) and (7) except for Q_*, is a function of the state of the system, importantly the vertical distribution of temperature, T_0, the composition of the atmosphere and the surface properties. E_c is the effective emitted irradiance of cloudy scenes that is calculated by compositing over all columns not identified as being cloud free. It depends on the distribution of clouds, but because clouds have some transparency and the atmosphere has some opacity, it also depends on the composition and temperature of the atmosphere, and the co-variability among the two. The short- and longwave components of the CRE, denoted by subscripts (Q) and (E) respectively, follow naturally as the difference between the all-sky radiative flux and the fluxes which would be manifest in the absence of clouds, *i.e.,*

$$F_\mathrm{c}^{(Q)} = -Q_*(\alpha_\mathrm{c} - \alpha_0) A_\mathrm{c} \tag{8}$$
$$F_\mathrm{c}^{(E)} = -(E_\mathrm{c} - E_0) A_\mathrm{c}. \tag{9}$$

[3] By conservation of energy, and as other sources of energy in the Earth system are negligible [Pilewskie *et al.,* this issue] this time-rate of change in Earth's enthalpy is equal to the net flux at the top of the atmosphere, N. Because the atmospheric heat content changes so little, a consequence of the low heat capacity of the atmosphere, N at the top of the atmosphere is usually taken as identical with N at the surface. The enthalpy notation is adopted because it emphasizes what is being measured, namely the change in the heat content, or enthalpy of the Earth system.

Given our sign convention, and because α_c is generally greater than α_0 whereas E_c is generally less than E_0, the shortwave CRE is negative and the longwave CRE is positive. Both quantities increase in magnitude with cloud amount, A_c. The net CRE, F_c is given by $F_c^{(Q)} + F_c^{(E)}$. A secular increase in the magnitude of $F_c^{(Q)}$ would exert a cooling influence on the Earth system, whereas an increase in the magnitude of $F_c^{(E)}$ would exert a warming influence. It should be emphasized that the CRE depends not just on the properties of the cloudy fraction of the planet but on the *differences* between the cloudy and cloud-free portions of the planet. As reviewed by Loeb et al. (2009) the application of this concept to various data sets shows the shortwave CRE to range from -45.4 to -53.3 $\mathrm{W\,m^{-2}}$ and the longwave CRE to range from 26.5 to 30.6 $\mathrm{W\,m^{-2}}$; the associated net CRE from these prior estimates ranges from -16.7 to -24.5 $\mathrm{W\,m^{-2}}$. The CERES EBAF data, upon which the TOA estimates in Fig. 1 are largely based, give a shortwave cloud radiative effect of -47.1 $\mathrm{W\,m^{-2}}$ and a longwave CRE of 26.5 $\mathrm{W\,m^{-2}}$; overall clouds, more precisely cloudy-scenes, exert a net cooling influence on the Earth system of about 20 $\mathrm{W\,m^{-2}}$.

The CRE concept has seen considerable use in the interpretation of feedbacks in the climate system. From the perspective of Eq. (1) a feedback is a change in a radiative flux that results from a change in global temperature; such a further change in radiative flux, in addition to that caused by an initial forcing imposed on the climate system, can enhance or diminish the temperature change induced by a given forcing (positive or negative feedback, respectively). The feedback concept has been quite useful in interpreting the contributions of different components of the climate system to S_{eq}. It is quite straightforward to demonstrate that changes in cloud radiative effects are not the same as cloud feedbacks. To appreciate this point note that, in the limit of small changes, the strength of the shortwave cloud feedback can be derived formally from (6) and (7) as (*e.g.,* Soden et al., 2008; Schwartz, 2011)

$$\lambda_c^{(Q)} = \frac{\partial Q}{\partial A_c}\frac{\partial A_c}{\partial T_s} + \frac{\partial Q}{\partial \alpha_c}\frac{\partial \alpha_c}{\partial T_s} \tag{10}$$

$$= -Q_* \left[(\alpha_c - \alpha_0)\frac{\delta A_c}{\delta T_s} + A_c \frac{\delta \alpha_c}{\delta T_s} \right]. \tag{11}$$

A change in the CRE from surface temperature induced changes in cloud amount is, however, not simply equal to the cloud feedback times the change in surface temperature; but rather includes an additional term that accounts for the change in the albedo of the cloud-free scenes with the change in surface temperature:

$$\delta F_c^{(Q)} = \lambda_c^{(Q)} \delta T_s - Q_* A_c \frac{\partial \alpha_0}{\partial T_s} \delta T_s. \tag{12}$$

The relationship between the cloud feedback and the change in the CRE follows similarly for the longwave part of the spectrum. Eq.(12) explicitly includes the dependence of the CRE on factors other than cloudiness, in the present example also the cloud-free sky albedo and emission. Thus if, in a changing climate, the surface albedo or cloud-free-sky aerosol changed, the CRE, as conventionally defined, would change for reasons that have nothing to do with changes in cloud properties. If the cloud properties remained fixed, the actual cloud feedback would be zero. This situation is described by Soden et al. (2008) as a masking effect. This example is readily extended to the longwave, where the change in the cloud-free-sky emission with the further caveat that the difference between changes in CRE and cloud feedbacks depends on how the problem is conceptualized, *i.e.,* how Eqs. (6)-(7) are formulated to begin with. Soden et al. (2008) illustrate how the use of radiative kernels can be used to diagnose cloud feedbacks from changes in CRE, although this method conflates cloud feedbacks with cloud mediated CO_2 indirect forcings, as discussed below. Because δE_0 and $\delta \alpha_0$ are in principal observable, the net effect of changing clouds on the response of the system to an external perturbation can, in principle, be determined from measurements.

A further concern with the cloud forcing concept is that the CRE, as defined above, depends not only on the properties of the cloud-free scenes, but also on other radiation-influencing constituents of the climate system that might be correlated with the presence of clouds. Hence it is really a *cloudy-scene* radiative effect, *i.e.,* the difference between the total radiative flux, and that which would be measured if cloud-free conditions always prevailed. The emphasis on cloudy-scenes rather than clouds is because the former admits the possibility that the atmosphere in cloudy scenes is systematically different from the atmosphere in cloud-free scenes. For instance, if the atmosphere in cloudy scenes tends to be more humid than in cloud-free scenes then this difference in humidity will, through the definition of cloud radiative effect (9), be interpreted as an effect of clouds. This distinction has important consequences for how cloud radiative effects are calculated in models versus in measurements. In models the radiative transfer calculation is usually performed twice, the second time with the clouds removed from the input. The first call to the radiation defines the all-sky radiative flux, the second call defines the clear-sky radiative flux and their difference is called the cloud radiative effect. In observations one cannot usually remove clouds,[4] so the clear-sky radiative flux is estimated based on scenes

[4] Although there are exceptions as Intrieri et al. (2002) and Mauritsen et al. (2011) calculate the clear-sky radiative fluxes using measurements of the atmosphere in cloudy conditions, but do not include the clouds in the radiative transfer calculations. This however is not a measurement of radiative effects, rather a calculation of radiative effects given measured atmospheric properties.

where no clouds are identified in the first place, rather than based on all scenes but with removal only of the clouds from those scenes in which they occur. If the atmosphere around the clouds is different (for instance more humid) in the cloudy scenes, as one might expect, this leads to differences between the two estimates of CRE. Such effects were shown by Sohn et al. (2010) to lead to systematic discrepancies between the observed longwave CRE and that calculated by models. This finding calls into question the current practice of calculating CRE in models through double calls to the radiative transfer and all the more that of calculating a change in CRE.

It does not appear to be particularly helpful to conceptualize clouds as an external component of the climate system, as is routinely done for quantities such as the solar constant or the anthropogenic contribution to the greenhouse gas concentration, and as such there appears to be little need to speak of cloud "forcing". However, if clouds are considered as an internal component of the system they can contribute to the forcing associated with externally imposed changes in other constituents. For example, it is usually assumed that only the longwave emission of the atmosphere depends on the greenhouse gas concentration, so that $E = E(\chi, T_s)$ where χ denotes the greenhouse gas concentration. In this case the net radiative forcing that results from a perturbation in a greenhouse gas concentration, $\delta\chi$, can be expressed as follows

$$F_\chi = F_\chi^{(E)} = -\frac{\partial E}{\partial \chi}\delta\chi. \qquad (13)$$

To the extent that cloud amount also depends on the greenhouse gas concentrations there is a resulting further contribution to the radiative forcing of the greenhouse gas perturbation, so that the total radiative forcing due to the perturbation in concentration becomes

$$F_\chi = \left[\frac{\partial Q}{\partial A_c}\frac{\partial A_c}{\partial \chi} - \frac{\partial E}{\partial \chi}\right]\delta\chi. \qquad (14)$$

The first term on the right-hand side introduces the idea of an indirect forcing of greenhouse gases that is mediated by clouds. The use of the word "indirect" signifies that the change in the shortwave flux is not a direct consequence of the greenhouse gas concentration in the way the change in atmosphere emissivity is, but rather results from the sensitivity of cloud amount to the longwave emissivity of the atmosphere. Given our description of the system through Eqs. (6) and (7), an indirect CO_2 forcing follows as soon as one admits that cloudiness may depend on the concentration of atmospheric CO_2 (Forster and Gregory, 2006; Gregory and Webb, 2008; Andrews et al., 2009). The idea that clouds may be sensitive to the concentration of greenhouse gases, and thus act as an indirect forcing, actually predates the idea that clouds may depend on surface temperature and hence act as a feedback (Plass, 1956). The physical mechanism is that the radiative cooling at the top of stratiform cloud layers, which is important to their sustenance, is sensitive to the downwelling longwave radiative flux, which in turn depends on the longwave opacity of the overlying atmosphere (Caldwell and Bretherton, 2009; Stevens et al., 2003).

3.2 Cloud amount

A more intrinsic concern regarding quantification of the influence of clouds on the climate system is that the utility of many of the above ideas is premised on the idea that a cloud is a well defined entity that can be identified without ambiguity. However this is far from the actual situation. Clouds are, in a word, nebulous. A cloud, like an aerosol more generally, is a dispersion of particulate matter in a turbulent flow. But, additionally and in contrast to clear-air aerosols, clouds are inherently ephemeral, as they contain a substantial amount of condensed (liquid or solid) water, the presence and amount of which are maintained by local supersaturation, and which can quickly dissipate by evaporation, converting a cloudy scene to a cloud-free scene, making it difficult to determine the boundaries or even the presence of a cloud. This situation leads inevitably to a certain arbitrariness in whether a cloud is present at a given location and thus in what constitutes cloudy or cloud-free scenes.

It might be argued that a clear basis for defining a cloud is provided by Köhler theory, namely as the set of particles that exist in an environment that is supersaturated relative to the equilibrium supersaturation over the particle surface and for which the equilibrium state is unstable. However, such a definition is not very useful in practice as the theory applies at best only to liquid clouds, and deliquesced aerosol in humid environments, or evaporating hydrometers in subsaturated environments are often optically indistinguishable from clouds defined on the basis of Köhler theory, the transition region extending over distances up to kilometers or more (Koren et al., 2007; Tackett and Di Girolamo, 2009; Twohy et al., 2009; Bar-Or et al., 2010).

As a consequence of such concerns the presence of clouds is often determined based on their radiative properties. This approach is certainly advantageous from practical perspectives, especially as satellites can be built to be sensitive to such properties and thus afford the opportunity for reproducible measurements with global coverage and high spatial and temporal resolution. Whether this approach is suitable remains an open question, the answer to which depends on the use being made of the observations. The real question would thus seem to be whether the solutions provide a product that is fit for the intended use. One important such use is to serve as a basis in measurement for evaluation of climate models

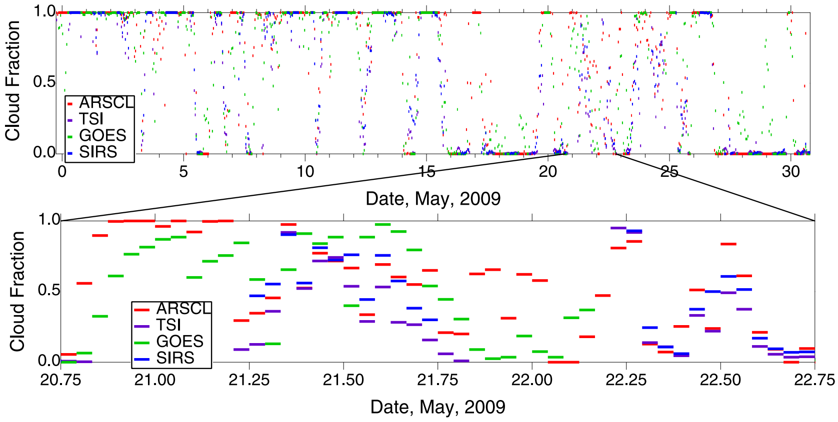

Fig. 3 Cloud Fraction in north central Oklahoma determined by several techniques. Both panels show 3 hr mean cloud fraction as a function of local standard time; upper panel gives one-month overview; lower panel expands a 2-day period. ARSCL (Active Remote Sensing of Clouds) product of the Department of Energy Atmospheric Radiation Measurement (ARM) Program (Clothiaux et al., 2000), is time-average based on vertically pointing lidars and millimeter cloud radars; SIRS (Solar Infrared Radiation System) product of the ARM Program Long et al. (2006) is time-average based on downwelling shortwave irradiance within nominal 160 field of view. TSI (Total Sky Imager) is based on fraction of cloudy pixels within hemispheric field of view GOES (Geostationary Operational Environmental Satellite) is based on average of all pixels (4 km pixel size) within 20 km of the surface measurement site (Genkova et al., 2004). Data provided by W. Wu (Brookhaven National Laboratory, 2011).

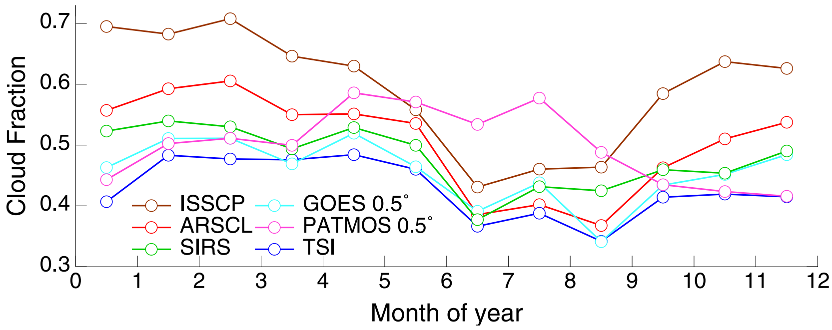

Fig. 4 Monthly mean cloud fraction in north central Oklahoma determined by several techniques, as in Fig. 3. ISSCP reference imaging radiometers on operational weather satellites (ISCCP); PATMOS-X Pathfinder Atmospheres Extended Product (Heidinger and Pavolonis, 2009). Data provided by W. Wu (Brookhaven National Laboratory, 2011).

and for interpretation of cloud radiative effects and feedbacks all of which rely on measurement accuracy. A second important use is for the examination of temporal trends, which requires reproducibility and precision. One approach to assessing measurement accuracy is by comparison of frequency distributions. Rossow and Cairns (1995) show similar U-shaped frequency distributions of cloud fraction between satellite observations and visual surface observations, albeit with substantial differences, especially at the high-cloud-cover end of the distribution (surface observations higher than satellite). A more stringent test would seem to be comparison of time series of measurements that are co-located in space and time. As shown in Figs 3-5 such co-located measurements by multiple approaches yield results that differ so greatly in measurements at a single time, in monthly averages, in the seasonal pattern, and even in monthly anomalies that they call into question the utility of the quantity cloud fraction. The sensitivity of cloud fraction to threshold and resolution has long been recognized (*e.g.,* Wielicki and Welch, 1986; Di Girolamo and Davies, 1997; Rossow and Schiffer, 1999; Foster et al., 2010), and various solutions to these problems have been proposed. In this respect it would seem that the central problem with cloud fraction is that, as a dispersion, clouds do not naturally admit well-defined boundaries, a problem which is compounded by the turbulent nature of the medium in which clouds are found. For this reason, bulk quantities such as cloud optical depth, liquid- or solid-water path, or radar reflectivity, can be expected to vary continuously and thus be less sensitive to threshold and resolution, and thereby provide a more useful characterization of cloudiness.

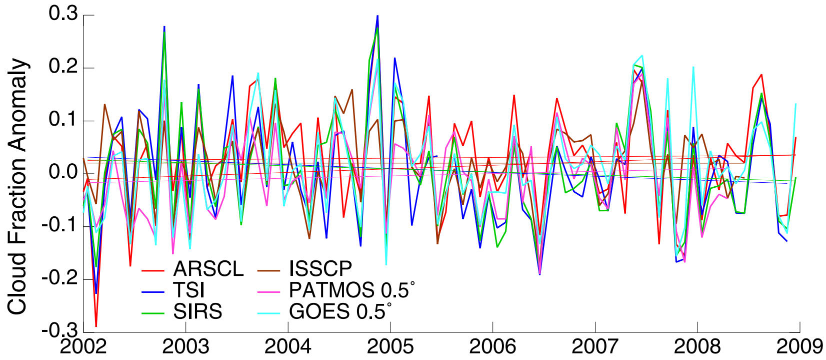

Fig. 5 Anomalies in monthly mean cloud fraction in north Central Oklahoma determined by several techniques, as in Fig. 3. Trend lines are all zero within approximately one sigma, as evaluated assuming the absence of autocorrelation. Data provided by W. Wu (Brookhaven National Laboratory, 2011).

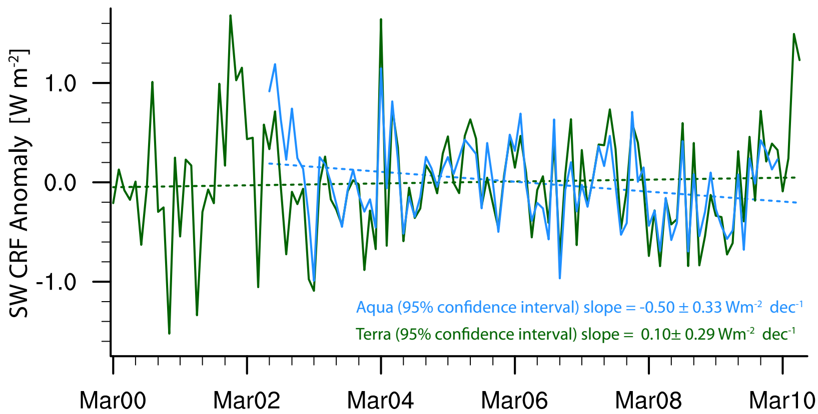

Fig. 6 Anomalies in the monthly and globally averaged top-of-atmosphere shortwave cloud radiative effect (CRE) from the CERES SSF1deg product. Cloud shortwave forcing, difference between all-sky and cloud-free downwelling fluxes at the TOA, *e.g.,* Eq. (8), is negative. Anomalies are calculated by subtracting the CERES monthly mean cloud forcings (average over the data set) from the individual monthly forcings. Negative slope, *i.e.,* decrease in cloud forcing with time, corresponds to increase in cloud reflectivity climatology from the monthly fluxes.

3.3 Trends in cloudiness

Assessment of temporal trends is best accomplished through examination of anomalies (departure from average over period of record). This approach removes seasonal variation (which is much greater than the trend being sought) and also removes biases between different approaches, permitting comparison of trends across approaches, while preserving measurement precision. Some of the consequences that arise from both issues of definition and differences in observational techniques for cloud fraction are illustrated in Fig. 5 through a comparison of monthly mean cloud fraction anomaly derived from different products. No significant trend is indicated. An examination (Foster et al., 2009) of global mean monthly cloud amount anomalies 1971-2008 from a number of sources likewise indicated that the variations and their differences were smaller than the estimated uncertainties and much smaller than the annual cycle leading to the conclusion that the global monthly mean cloud cover has changed by less than a few percent over more than two decades. Recognition of the importance of clouds as agents of feedback in the climate system and the changes in forcing over that period of time prompted the observation by Foster et al. (2009) that “the lack of change [in cloudiness] needs to be explained as much as would a significant change.” This of course begs the question as to what constitutes a substantial change.

From the perspective of climate change, a change in cloud irradiance that is comparable to a given radiative forcing over the time period of interest can be considered substantial. Over the period 1960-2005 CO_2 increased at an average rate of 1.4 ppm yr^{-1} (Forster et al., 2007). A 3.7 $W\,m^{-2}$ forcing associated with a doubling of CO_2 Forster et al. (2007) would result in a trend in the energy budget of about 0.25 $W\,m^{-2}\,dec^{-1}$, in the absence of any feedbacks. In this context a change in cloud properties in response to this forcing that resulted in a further change in the radiation budget $|\dot{Q}| > 0.05$ $W\,m^{-2}\,dec^{-1}$ would constitute an appreciable change. Based on this we argue that observationally constraining the response of the climate system to such a perturbation to some meaningful degree requires an ability to detect changes in the radiation budget $|\dot{Q}|$ of 0.05 $\mathrm{W\,m^{-2}\,dec^{-1}}$. The analysis of Loeb et al. (2007) suggests detection of such a trend with 90% confidence would require fifty years of data, and this analysis is optimistic as it assumes a perfect instrument and insignificant decadal variability. Moreover, as the number of years of data required to establish a trend is proportional to $|\dot{Q}|^{-2/3}$, where $|\dot{Q}|$ is the magnitude of the trend (Weatherhead et al., 2000), establishing a trend even twice as large would still require thirty years.

Looking at a decade of CERES SSF1deg shortwave CRE data suggests that even these estimates are optimistic, as over the past decade data from the CERES instruments on the Terra and Aqua satellites exhibit different trends. This is illustrated with the help of Fig. 6 which shows the monthly anomalies from each platform, and the range of decadal trends at the 95% significance level. That the 95% significance ranges of the trends in monthly global anomalies barely overlap suggests either that the trends in Q_c changes through the course of the diurnal cycle, and/or that there are unaccounted for stability problems in the instruments. Barring a more sophisticated analysis it is difficult to constrain $\dot{Q}_c$ to within $\pm 1\ \mathrm{W\,m^{-2}\,dec^{-1}}$. It would seem that this situation would persist unless and until the reasons for the difference in these trend lines is determined, and thus that any determination of change in CRE and, all the more, the relation of such change to either the changing composition of the system, or the accompanying changes in surface temperature, must be held in abeyance.

Because changes in the measured CRE incorporate a variety of effects, ranging from cloud feedbacks to CO_2 indirect forcings to changes in other components of the system that may correlate with clouds, the measurements require auxiliary hypotheses if they are to be used to constrain the models. Nonetheless, a practical step toward a better interpretation of the measurements would be to begin calculating the CRE in the models in a way that mimics the way this is calculated from the data. Such a step will increasingly be facilitated by the move to higher resolution in climate models, with the resultant increase in likelihood of finding cloud free columns. One possible approach that may prove useful is examination of changes in cloudiness at regional scales, rather than only at global scales. Regional variation in low-level clouds over the Northeast Pacific has been found to be inversely correlated with variations in sea surface temperature (Clement et al., 2009) leading the investigators to suggest that clouds act as a positive feedback in this region on decadal time scales. Still enthusiasm for such an approach must be tempered by the realization that it cannot distinguish changes in cloudiness that are in response to changes in globally averaged surface temperature from changes that would arise, for example, from changes in regional aerosol loadings or regional circulation changes unrelated to a global temperature increase.

4 Aerosol radiative effects and forcing

4.1 Aerosol radiative effects

The radiative properties of cloud-free skies depend in addition to surface reflectivity, on the gaseous composition of the atmosphere and on the spatial and temporal distribution and chemical and microphysical properties of atmospheric aerosols. Whereas aerosol effects on the shortwave radiative properties of the atmosphere are usually emphasized (Charlson et al., 1992), aerosol effects on the emissivity of the atmosphere may also be important (Markowicz et al., 2003). Atmospheric aerosol particles also serve as cloud condensation nuclei CCN and ice forming nuclei IFN that are the seed particles for formation of cloud liquid drops and ice crystals, and thus changes in the aerosol are expected to change cloud properties, thereby contributing to an indirect aerosol radiative forcing. (Twomey, 1974; Charlson et al., 1992; Albrecht, 1989; Stevens and Feingold, 2009). Hence atmospheric aerosols influence, to varying degrees, the flows of energy in almost all of the constituent parts of Fig. 1. Because human activity has long been recognized as contributing to increasing aerosol burdens, aerosols also contribute to the net radiative forcing that appears in Eq. (4). The radiative forcing of aerosol is generally appreciated to be negative. As a consequence taking this forcing into account in estimating climate sensitivity Eq. (4) results in a greater sensitivity than would be inferred from consideration only of the greenhouse gas forcing (Gregory et al., 2002; Schwartz et al., 2010). This situation necessitates accurate knowledge of the aerosol forcing, which in turn requires understanding how the aerosol burden has changed as a result of human activity and translating this changing aerosol burden into a radiative forcing. However this understanding is poor, and hence the the magnitude of the aerosol radiative forcing is highly uncertain (Forster et al., 2007).

This poor understanding is a direct result of the complexity of aerosols and their radiative influences, a consequence both of their heterogeneous chemical and microphysical properties and also their highly variable spatial and temporal distribution. The heterogeneous composition, which contrasts with the well defined molecular properties of the greenhouse gases, is a consequence of the numerous contributions to atmospheric aerosols: Primary emissions from natural and anthropogenic sources and gas-to-particle conversion resulting from atmospheric reactions of precursor gases, importantly sulfur and nitrogen oxides from combustion sources, ammonia from agriculture and animal husbandry, organics from anthropogenic sources and vegetation, and numerous other sources. Gas-to-particle conversion processes lead

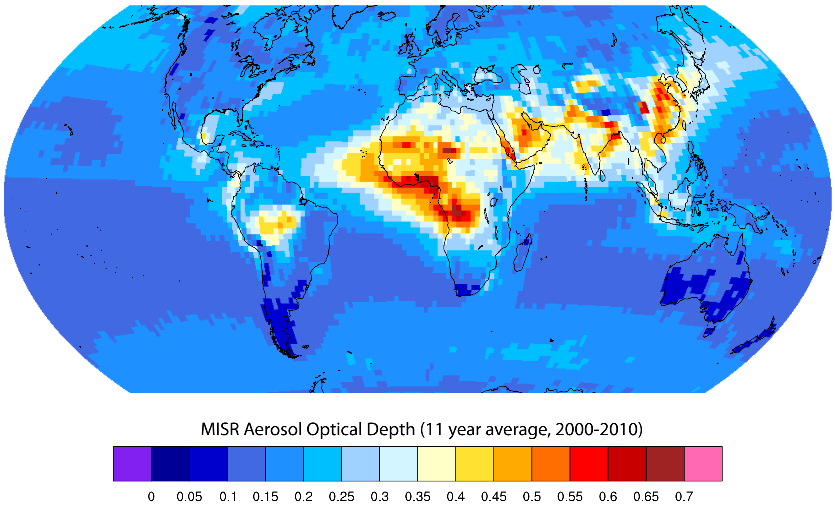

Fig. 7 Annually averaged aerosol optical depth climatology inferred from eleven (2000-2010) years of MISR (Multi-angle Imaging Spectro Radiometer) measurements.

both to new particle formation and to growth of pre-existing particles. The resulting aerosols undergo further evolution in the atmosphere through condensation and coagulation and in-cloud processing. Ultimately the aerosol particles are removed from the atmosphere, importantly by precipitation. The optical and cloud-nucleating properties of aerosols, and thus their influences on climate and climate change, are strongly dependent on the size and chemical composition of the particles comprising the aerosol. For example growth of particles with increasing relative humidity, which greatly increases their ability to scatter visible light, is highly dependent on composition.

The complexity in the spatio-temporal distribution of aerosols is hinted at even upon inspection of their long-term average global distribution undifferentiated by aerosol type as seen in the large spatial variability of aerosol optical depth (AOD) averaged over the available eleven years of Multi-angle Imaging SpectroRadiometer (MISR) data, Figure 7. Major contributions arise from windblown dust (*e.g.,* Northern Africa, Arabian peninsula, western China) and biomass burning (*e.g.,* central Africa, Amazonia, Indonesia). Substantial contributions from human activity, mainly combustion related, are evident in southeast and eastern Asia, extending into the western North Pacific. The highly industrialized regions of Europe and North America (extending to the North Atlantic) also exhibit noticeable enhancement of AOD relative to pristine continental regions and major portions of the Southern Hemisphere Ocean. The spatial heterogeneity of the distribution of these aerosols is a consequence of the heterogeneous distribution of sources together with the short atmospheric residence times of these aerosols, about a week, together with the intermittent removal by precipitation. Because aerosol sources have pronounced seasonality, and because sink and transport processes of all aerosols are heavily dependent on variable meteorological conditions, the distribution of aerosols shown in Fig. 7, being a long-term average, considerably understates the complexity of the spatial distribution of atmospheric aerosols. This points out that the complexity of aerosols is manifested not only by their varied chemical and microphysical properties, but also by their heterogeneous spatial distribution; for example an aerosol particle above a bright surface has a different radiative effect compared to even the same particle over a darker surface. All of which makes it much more difficult to quantify aerosol forcing than is the case with the incremental greenhouse gases.

The short residence time of aerosol particles in the troposphere not only complicates characterization of their radiative influences, but also has implications on climate change that would result from future changes in emissions, especially as most of the incremental aerosol arises from emissions associated with fossil fuel combustion. If at some point in the future emissions of CO_2 from combustion are substantially reduced, in response to recognition of their warming influence on climate, and if this were accompanied by reduction of associated emissions of sulfur and nitrogen oxides, major precursors of light-scattering tropospheric aerosols, the result would likely be for temperatures to initially increase because of the reduction of aerosol forcing. An initially abrupt increase in temperature following an abrupt cessation of aerosol forcing has been shown in climate model studies (*e.g.* Brasseur and Roeckner, 2005; Matthews and Caldeira, 2007).

As a consequence of all these considerations, understanding of energy flows in the Earth system and changes in these flows over the industrial era is challenged by poor understanding of the effect of aerosols on cloud-free skies where one can speak of the direct aerosol radiative forcing. This challenge is even greater with respect to aerosol effects on clouds (Lohmann and Feichter, 2005; Stevens and Feingold, 2009) where it has become common to speak of the indirect aerosol radiative forcing that results from modification of the radiative influences of clouds resulting from changes in cloud microphysical properties by anthropogenic aerosols. The discrimination of aerosol forcing into direct and indirect components structures thinking about aerosol influences on climate and is used to structure the discussion below, although it merits emphasizing that the utility of such a concept remains questionable. Except in special cases, there is little evidence that radiative forcings arising from different constituent perturbations are additive, so that a negative aerosol forcing cannot be expected to exactly compensate a positive greenhouse gas forcing of the same magnitude, even for simple metrics such as the globally averaged surface temperature (*e.g.,* Hansen et al., 1997; Shindell and Faluvegi, 2009).

4.2 Aerosol direct forcing

As a first step in estimating the direct aerosol radiative forcing it is necessary to quantify the magnitude of aerosol loading. The magnitude of aerosol loading is commonly expressed by the aerosol optical depth, AOD, the vertical integral of the aerosol extinction coefficient, typically given in the mid-visible, 500 or 550 nm. This quantity can, with diligence, be measured by sun photometry during the daytime and in the absence of clouds in the path to the sun to an accuracy of 0.01 (*e.g.,* Holben et al., 2001; Kim et al., 2008; Michalsky et al., 2010); the aerosol contribution to path extinction is determined by subtraction of extinction due to Rayleigh scattering and to atmospheric gases (importantly ozone). However such measurements are limited in their spatial coverage and to cloud-free skies during the day and are especially lacking over the oceans (Smirnov et al., 2009). Hence the approach taken to characterizing the distribution of aerosols globally has been to determine AOD by satellite, as for example illustrated by Fig. 7. It should be stressed, however, that the satellite measurement of AOD is rather indirect. The AOD is inferred from the enhancement of path radiance over that which would be obtained in the absence of the aerosol: Rayleigh scattering and surface-leaving radiance. Rigorous cloud screening, which recalls the discussion (§3.2) of what is a cloud, is also required. After contributions to path radiance from the surface and from Rayleigh scattering are subtracted the remaining contribution to path radiance is attributed to light scattering by the aerosol. Converting from radiance to optical depth rests on assumptions about the scattering phase function (angular distribution of light scattering) and the fraction of the extinction by the aerosol that is due to scattering versus absorption. Typically look-up tables are used, with identification of aerosol type informed by the wavelength dependence of the aerosol light scattering and climatology of aerosol types. The MISR (Multi-angle Imaging SpectroRadiometer) instrument, upon which the measurements illustrated in Fig. 7 are based, takes advantage of measurements at multiple scattering angles to better constrain the phase function and in turn the AOD [Kahn, this issue]. This is particularly helpful over bright and complex surfaces, because the inferred AOD is highly sensitive to errors in estimates of surface-leaving radiance. Measurements at multiple angles also constrains inferences of aerosol type, and hence absorption, which is another large source of error in inferring AOD. Further detail on determination of AOD by satellite and associated uncertainties is provided by Kahn et al. (2010) – see also Kahn, this issue.

Although the aerosol optical depth is the more commonly reported aerosol extensive property, the quantity of principal interest in the context of understanding the direct radiative influence of aerosols on climate is the forcing, the change in irradiance at the TOA. Here some care must be given to the definition of direct aerosol forcing. Analogous to the way in which a CRE is calculated, the direct aerosol radiative effect, DARE, is the change in the irradiance at the top of the atmosphere that results from the total aerosol present. For purposes of quantitative comparison with other forcings such as greenhouse gas forcings, the DARE and other measures of aerosol forcing are generally presented as 24-hour averages. The principal contribution to DARE is in the shortwave, and it is this quantity that is generally reported. The quantity pertinent to forcing of climate change, direct aerosol radiative forcing, DARF, is as is the case with greenhouse gas forcing, not the entire aerosol radiative effect but the secular (mainly anthropogenic) change in the DARE due to incremental aerosols, as that is the externally forced contribution. A negative forcing denotes a decrease in absorbed shortwave irradiance; opposite in sign to the positive greenhouse gas forcing, and, within the framework of the forcing-response paradigm, would offset some fraction of the greenhouse gas forcing.

A further consideration pertinent to determining DARE or DARF is that just as a part of the difference in irradiance between all sky and cloud-free scenes that is attributed to the CRE can result from differences in the cloud environment, as distinct from changes in the amount or properties of clouds, part of what might be attributed to the DARE might likewise result from differences in atmospheric composition or surface reflectance that covary with aerosol loading. Natural variability in circulations, which for instance might result in more or less dust transport over the Atlantic would, in this framework, be interpreted as a change in the DARE, and hence an aerosol radiative forcing, even if its cause had nothing to do with the aerosol per se.

Determining the DARF requires knowledge of the radiative effects of preindustrial aerosols, as this is the background state about which a perturbation is defined. In principle this can be accounted for in models by modifying the source strengths of aerosols and aerosol precursor gases. Observationally the approach is to attribute accumulation mode aerosols, aerosols which are formed by gas-to particle conversion, and which operationally have diameter less than about

1 μm, to the anthropogenic perturbation, as the principal contributions to natural aerosol, sea spray and mineral dust, have particle diameters greater than 1 μm. Still this approach has its limitations; for example it would misattribute dust aerosol from tilled soils and likewise would misattribute natural organic haze such as gave rise to the name of the Great Smoky Mountains in the eastern United States.

Importantly, in determining the DARE, and how it has changed, it is necessary to account for the effects of clouds. The effect of cloud contamination of pixels used to determine aerosol optical depth as already been noted; such contamination would result in a gross overestimation of aerosol optical depth, and it is thus necessary to apply stringent cloud screening (Mishchenko et al., 1999). This has the effect of limiting determination of aerosol optical depth to scenes where relative humidity is lower than average, with the resultant effect that the swelling of particles at high relative humidity and the attendant increase in light scattering cross-section, optical depth, and forcing is excluded from the measurements. This effect can be substantial; for sulfate aerosols the scattering cross-section increases fourfold between 90 and 97% relative humidity (Nemesure et al., 1995). Surface-based measurements are also subject to a similar concern as the technique requires a direct path to the sun. More intrinsic to the radiative forcing issue, the direct radiative effect of light scattering by aerosols is greatly diminished in the presence of clouds, which prevents solar irradiance from reaching the aerosol if clouds are above the aerosol, or by providing a bright underlying albedo, minimizing the effect of aerosol scattering if clouds are below the aerosol. Clouds beneath an absorbing aerosol greatly increase the amount of solar absorbed irradiance relative to the cloud-free situation. Because it is difficult to retrieve aerosol amounts in the presence of clouds, it is has been common to estimate the DARE on a global basis simply by multiplying the cloud-free DARE, determined by measurement or by modeling of the amount and optical properties of the aerosol, by the cloud-free sky fraction. Such an approach assumes has been shown in model calculations to yield a value of DARF that is erroneously large (negative) by a factor of two (Bellouin et al., 2008) .

Increases in the amount of aerosol loading that can confidently be ascribed to anthropogenic emissions have been thought to give rise to changes in global average atmospheric radiative fluxes (aerosol forcings) that are a substantial fraction of the greenhouse gas forcings over the industrial era. As part of the fourth assessment report of the IPCC the DARF was estimated to be -0.5 ± 0.4 $\mathrm{W\,m^{-2}}$ (Forster et al., 2007, 90% confidence limits). A value of DARF at the high magnitude end of this range would offset a substantial fraction of greenhouse gas warming over the industrial era (about 3 $\mathrm{W\,m^{-2}}$); the lower total forcing, when used to infer climate sensitivity, would imply a substantially larger climate sensitivity than would be inferred from observations of temperature changes and radiative forcings from anthropogenic greenhouse gases alone. Estimates of such a large value of DARF have been obtained mainly from observational estimates which as noted above cannot properly account for cloud effects. Differences between models and observations are also large over land, where satellite based methods are also weakest. Further contribution to uncertainty in estimates of aerosol direct forcing arises from the sensitivity of the DARE to aerosol intensive properties. For a given amount of aerosol material, the forcing is dependent on the composition and size distribution of the aerosol (McComiskey et al., 2008), relative humidity, surface albedo, the presence of underlying or overlying clouds, and the like. Hence a better quantification of DARE is limited by an inability to fully characterize aerosol intensive properties, but also the nature of their covariability with clouds. Whereas research continues to focus on the former point, the latter is similarly important and largely overlooked.

In summary more recent research suggesting that the DARF is considerably less than earlier estimates, such as the value of -1.3 $\mathrm{W\,m^{-2}}$ initially estimated for the sulfate aerosol alone (Charlson et al., 1992) and toward the low end of the range given for this quantity by AR4. As these differences are substantial in the context of the fraction of greenhouse gas forcing that would be offset by aerosols, and would substantially reduce the uncertainty in aerosol forcing, resolving this would seem to be a fruitful area for near-term investigation. Because of the difficulty in characterizing aerosols in the presence of clouds, and especially above clouds, this is a particularly thorny problem. One approach that might prove useful is characterization of aerosol loading by lidar from space (Winker et al., 2007).

The issues of aerosol direct forcing can be examined from another perspective, by looking for signatures of aerosol trends in radiative fluxes measured by passive space borne instrumentation over the past decade of intensive earth observations. Eleven years of MISR measurements (upper panel of Fig. 8) of AOD show large-scale shifts in specific regions. Large increases in AOD over Southeast and East Asia are likely a consequence of industrialization in those regions. In the Middle East rapid development in Arab states of the Persian Gulf may contribute to the strong increase in AOD there. Over the region ranging from southwest of North America, across the Atlantic and into North Africa, the Mediterranean and central Europe, and over the maritime continent AODs have decreased sharply, by as much as 0.1. A modest decrease is evident across the southern ocean, and there is an apparent increase over western Canada. The uncertainties associated with the retrievals of AOD from space (Kahn et al., 2010), and the susceptibility of decadal trends to the effects of inter-annual modes of variability such as El Nino preclude drawing confident conclusions from Fig. 8; however, the broad conclusions drawn by the figure are also supported by a more systematic analysis based on multiple platforms (Zhang and Reid, 2010).

A concern over the interpretation of these measurements is that the marked changes in aerosol optical depth inferred from the MISR measurements are not mirrored in trends in the outgoing shortwave radiation in cloud-free scenes over

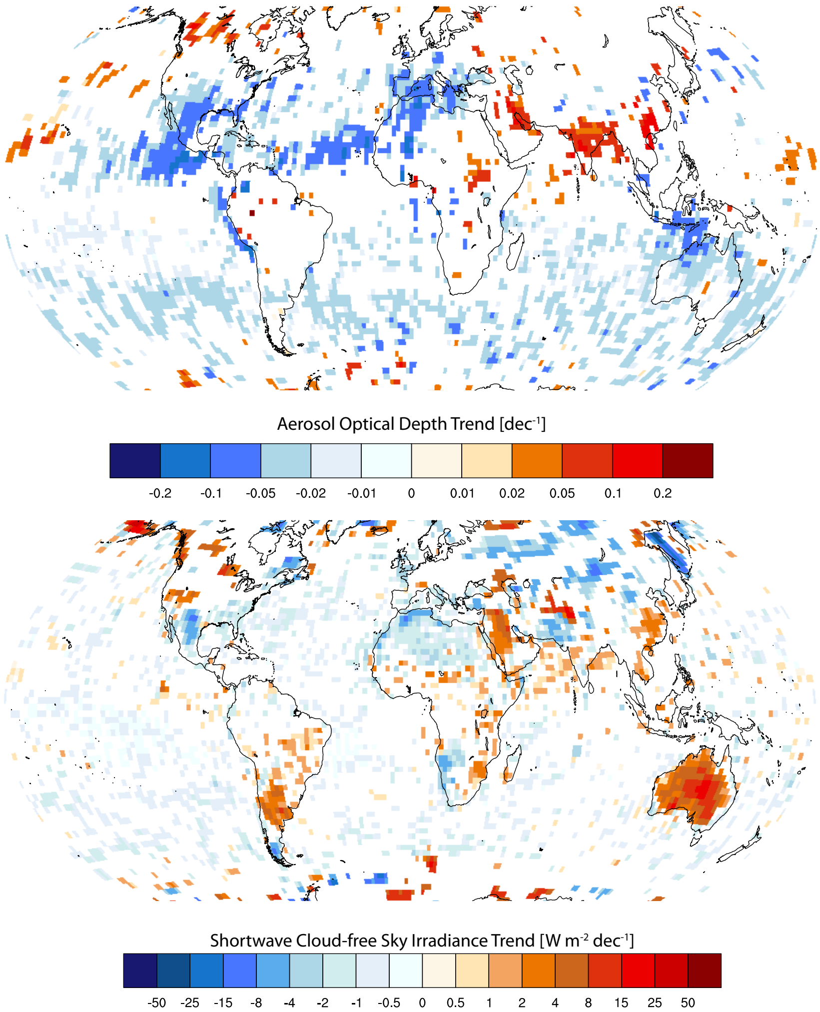

Fig. 8 Decadal trend in the annually averaged aerosol optical depth at (550nm) inferred from eleven years (2000-2010) of MISR (Multi-angle Imaging SpectroRadiometer) measurements (upper) and in cloud-free scene (TOA) radiative fluxes taken from CERES measurements (lower). Both the CERES and MISR instruments are aboard the Terra satellite. Only trends that are significantly different than zero at the 95% confidence level (as estimated based on the ratio of the residual variance and mean trend assuming no auto correlation in the yearly data) are shown. An increase in reflected shortwave irradiance, as indicated by the redder colors, is indicative of a brightening of the cloud-free skies.

the same time-period as measured by the highly calibrated CERES radiometer on the same platform, lower panel of Fig. 8. Although the radiative fluxes in cloud-free circumstances show trends in some regions where aerosol optical depth retrieved by MISR has been increasing, for instance over south Asia, particularly the middle east and south-east China, and to some extent over western Canada, the relationship between changes in AOD and changes in clear-sky shortwave radiative fluxes is not striking. The largest changes (irrespective of sign) in cloud-free-sky radiation appear over land in regions where there are no discernable trends in AOD. In almost every case statistically significant trends in CERES absorbed shortwave radiative irradiance over land are also evident in changes in land-surface properties measured over effectively the same time-period: the brightening, at TOA, of Australia; the brightening in Southeast Asia; the dimming of the Southwest US and darkening of Northern Mexico as well as the brightening/dimming patterns over Argentina and South America, and the dimming of the very northern tip of Africa are all associated with consistent changes in soil-moisture, whereby an increase in the cloud-free scenes outgoing shortwave radiation corresponds to a decrease in soil moisture and evapotranspiration (Jung et al., 2010).

The complexity of aerosols, especially the factors involved in their characterization and influencing their radiative forcing, raises the question as to whether it will be possible to reduce uncertainty in this forcing sufficiently to ultimately allow meaningful observational constraints to be placed on the equilibrium climate sensitivity. Certainly it would seem that the effort needed to reduce the uncertainty in aerosol forcing and its evolution over the industrial era globally must include an increased emphasis on cloud effects, and would dwarf the current research on aerosol forcing.

4.3 Aerosol indirect forcing

It has long been appreciated that tropospheric aerosols affect the microphysical properties of clouds affecting their albedo and precipitation development. Based on this understanding it has been hypothesized that changes in the loading and properties of tropospheric aerosols may indirectly affect the radiative influences of clouds (i.e., alter the CRE) by modifying cloud properties and/or amount. These changes are referred to as aerosol indirect effects, or the indirect radiative forcing resulting from anthropogenic (or secular) changes to tropospheric aerosol loading and properties. They are analogous to the CO_2 indirect forcings discussed in the context of Eq. (14). Because such indirect effects convolve changes in the aerosol with changes in cloudiness they tend to be complex and uncertain, and are only briefly touched upon them here.

Broadly speaking aerosol indirect effects can be classified into effects that result only from modification of the microphysical properties of clouds and those that result from ensuing changes to cloud macrophysical properties, for instance changes in cloud water content or cloud amount. Examples of the first type of effect, are hypotheses that an increase in the concentration of available cloud condensation nuclei results in a change in the number concentration (Twomey, 1974) and/or size distribution (Liu and Daum, 2002) of the drops in a cloud, and in turn in changes in the cloud radiative influence. Examples of the second type of effects include what have come to be known as cloud lifetime effects, such as hypothesized by Albrecht (1989) and Pincus and Baker (1994), wherein changes to cloud droplet number concentrations modify the precipitation efficiency of clouds, resulting in modifications of cloud amount (cf., Stevens and Feingold, 2009; Khain, 2009). Research continues to articulate hypotheses that fall into one of these two categories, many of which are beginning to to be extended to changes to ice-forming nuclei and hence the properties of ice or mixed phase clouds. Because of the magnitude of the CRE globally, a slight systematic change in the amount or albedo of clouds globally or over large regions of the Northern Hemisphere could exert a change in the global radiation budget that is comparable in magnitude to greenhouse gas forcing. For example, in a back-of-the-envelope calculation Charlson et al. (1992) showed that a 30% increase in the number concentration of cloud drops in marine stratus clouds globally would, other things being equal, exert a global radiative forcing of about -1 $\mathrm{W\,m^{-2}}$. In the intervening years there have been abundant climate model calculations of the global magnitude of aerosol indirect forcing over the industrial era, with estimates ranging from near zero to -3 $\mathrm{W\,m^{-2}}$ or more (Lohmann et al., 2010), based on differing assumptions regarding the relation between aerosols and cloud properties and various other controlling factors. In such a situation it would seem that the modeling is going far beyond the understanding.

To be sure, a wealth of observational support exists for Twomeys hypothesis that the cloud drop number concentration increases, cloud drop radius decreases, and cloud albedo increases with increasing aerosol particle concentration. However quantification of the net radiative forcing that can be attributable to the indirect effects of aerosols on clouds has proven elusive. That cloud drop radii are reduced and cloud albedo is increased by aerosols is clearly shown in ship tracks (*e.g.,* Segrin et al., 2007), but quantification of these influences is confounded by changes in cloud-water content. Regionally, reduction in cloud drop effective radius is associated with enhanced concentration of anthropogenic aerosol (*e.g.,* Schwartz et al., 2002), but the expected increase in cloud albedo is often absent. Such lack of enhancement of cloud albedo is likely due to a decrease in cloud liquid water path with increasing aerosol concentration, at variance with Twomey's *Ansatz* of other things, especially cloud water content, remaining equal. A global survey using satellite observations showed roughly equal likelihood of negative, near-zero, or positive correlation of column liquid water and column drop concentrations in liquid water clouds Han et al. (2002). Although several studies show strong correlations between cloud amount and aerosol optical depth (*e.g.,* Nakajima et al., 2001; Koren et al., 2010), the interpretation of such correlations is difficult, as a variety of processes (both physical and retrieval artifacts) can be expected to produce such correlations, quite independently of whether or not the aerosol is interacting with the cloud (*e.g.,* Loeb and Schuster, 2008). For example, both aerosol optical depth and cloudiness increase with humidity, and thus it is not surprising that modeling studies might overestimate aerosol indirect effects. By regressing the logarithm of the retrieved aerosol optical depth against the logarithm of the retrieved cloud droplet concentrations over a number of geographic regions, Quaas et al. (2005) found statistically significant slopes that range from 0.1 to 0.3 depending on location, with values over the ocean three times greater than those over land and with a global mean value of just under 0.2 (see also Quaas et al., 2009). Based on this analysis (Quaas et al., 2009) estimate quite a low aerosol indirect forcing, $0.2 \pm 0.1\ \mathrm{W\,m^{-2}}$. As noted by Quaas et al. (2005) this uncertainty is parametric, and contributions to the uncertainty from structural effects can be expected to be substantial; hence, observational estimates cannot, on their own, establish with confidence even the sign of the effect hypothesized by Twomey, despite arguments based on simple physical considerations that it is negative. The structural uncertainty that frustrates attempts to quantify the Twomey, or Twomey-like, effects also makes it difficult to test cloud lifetime hypotheses and all the more to quantify the resultant forcings. Moreover, to the extent that precipitation processes become involved the difficulties are compounded, in no small part because of the sensitivity of aerosol amount to wet scavenging by precipitation.

In sum, although there is little doubt of the importance of aerosol-cloud interactions in influencing the amount of atmospheric aerosol, the microphysical properties of clouds, and the macrophysical properties of clouds and precipitation development, the variety of ways in which clouds respond to microphysical perturbations and the tendency of cloud-scale circulations to buffer microphysical perturbations (Stevens and Feingold, 2009) lend weight to the argument that, after a full accounting, the radiative forcing attributable to aerosol effects on clouds is likely to be small, at least on a global scale.

5 Models

The above discussion serves as a reminder that the current observing system is not adequate to measure the energy flows through the climate system with sufficient accuracy to resolve some of the key challenges of climate science. This inadequacy reflects gaps in the observing system, but also the very nature of the Earth system, whose natural variability makes it difficult to detect trends on timescales as short as a decade. Hence it seems unavoidable that climate models will be necessary to advance understanding of energy flows in the climate system. Climate models are, however, as flawed as they are necessary. Although they can be useful to help understand observational data, they are not a replacement for these data, and likely will never be authoritative representations (in the sense that they provide the last word) of the climate system.

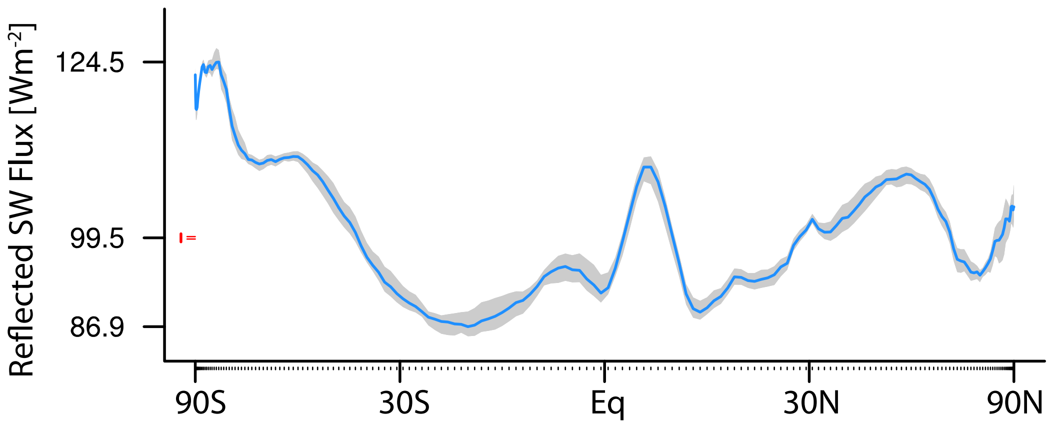

Fig. 9 Zonally averaged reflected shortwave radiation from ten years of CERES EBAF (energy balanced and filled) data. The figure shows $Q_n^{\uparrow}(\varphi)$, where φ denotes latitude, and the subscript n indexes the year. The year-to-year variability in $Q_n^{\uparrow}(\varphi)$ is indicated by the grey shading for each latitude. The blue line shows the ten-year mean as a function of latitude. The red vertical bar near the vertical axis denotes the range in the global and annual means. The two small horizontal bars near the vertical axis show the ten-year means averaged over the northern and southern hemisphere separately.

5.1 Two brief excursions in model evaluation

"All models are wrong; some are useful," (Box, 1979). This characterization of models certainly applies to climate models, which have proven themselves invaluable as a fertile methodology for developing and refining understanding of Earth's climate system. Indeed present climate models do a remarkably good job of reproducing many aspects of Earth's climate system albeit with significant departures from strict fidelity to the world as we know it. This point merits further exploration with the help of a couple of examples.

As a first example consider one of the key energy flows illustrated by Fig. 1, namely, the ability of the planet, despite the seemingly whimsical way in which the atmosphere distributes its condensate, somehow to maintain a surprisingly constant planetary albedo. This point holds true not only on globally and annually averaged scales, but even within latitude zones, as illustrated by the latitudinal dependence of zonal averages from a decade of CERES measurements of the reflected shortwave radiation at the TOA, Fig. 9. Also shown are the range of annual averages of the reflected shortwave radiation, evaluated as

$$\langle Q_n^{\uparrow} \rangle = \frac{1}{2} \int_{-\pi/2}^{\pi/2} Q_n^{\uparrow} \cos(\varphi) \mathrm{d}\varphi, \tag{15}$$

where $Q_n^{\uparrow}$ denotes the reflected shortwave radiation as a function of latitude φ and year n; and the values averaged over the northern and southern hemisphere separately. The surprising feature of Earth's climate system that is revealed in these measurements is the small inter-annual variability (the range in the yearly averages is 1.16 $\mathrm{W\,m^{-2}}$ and the standard deviation is 0.36 $\mathrm{W\,m^{-2}}$), despite the zonal average spanning more than 40 $\mathrm{W\,m^{-2}}$, with a root-mean squared variability (weighted by area) of nearly 9 $\mathrm{W\,m^{-2}}$. Also the difference between the two hemispherically averaged values is very small, only 0.35 $\mathrm{W\,m^{-2}}$.

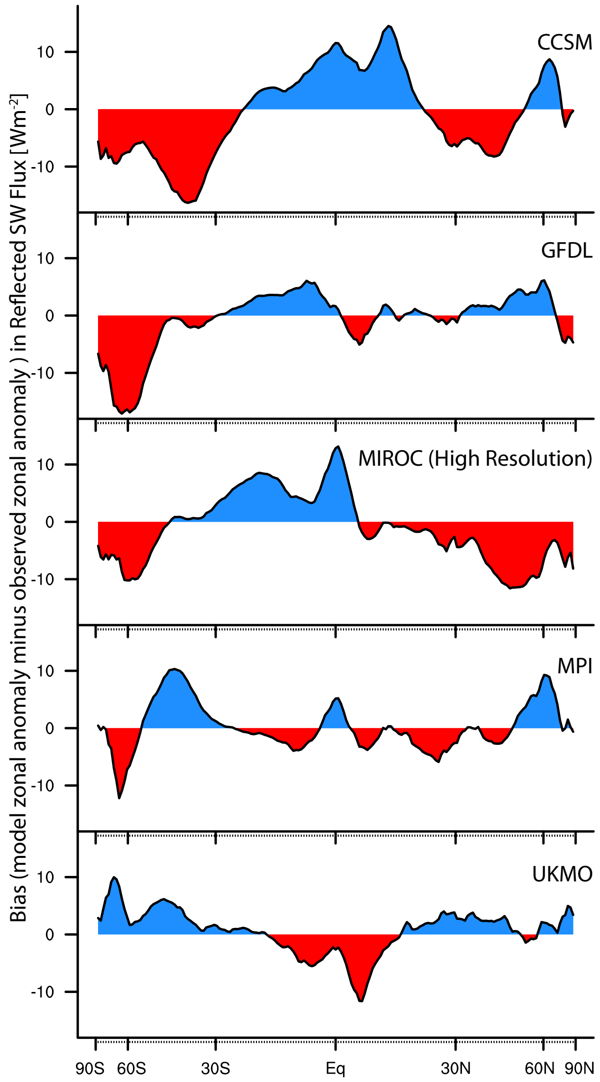

Fig. 10 Biases in zonal anomalies in annually and zonally averaged reflected shortwave radiation. Here the biases are the differences between modeled and observed estimates, so that a negative (red) anomaly corresponds to too much absorption in the model. The model estimates of $Q^{\uparrow\prime}$ are taken from five leading CMIP3 models, each processed over the last ten years of a slab-control simulation, the observations are taken from ten years of CERES data.

As the zonally averaged shortwave reflectance seems to be a robust property of the climate system, it is of interest to see how well it is reproduced in model calculations. As has been noted previously (*e.g.,* Bender et al., 2006), each of the models misrepresents important features of the planetary albedo, starting with the globally averaged value. Because the models have been adjusted to fit estimates of the global values, and at the time of CMIP3 the best estimates of the reflected solar radiation ranged from 101 to 106 $\mathrm{W\,m^{-2}}$; the difference between the simulated and observed global mean is not particularly meaningful. More pertinent as a basis for comparison between the models and observations are the departures from the zonal mean, Figure 10, as these zonal patterns help regulate the meridional heat transport and are not directly specified through the model development process. In this respect five leading climate models are characteristic of the borader set of models in that each exhibits some skill, with correlations between the observed and simulated latitudinal anomaly ranging between 0.60 and 0.94 for the models shown. However the departures from the observations, which are due mainly to treatment of clouds in the models, are substantial in the context of radiative forcings over the industrial era. The root-mean square error in the zonal residuals ranges from 4 to 8 $\mathrm{W\,m^{-2}}$, comparable to the departure of the zonal mean from the global mean and considerably greater than a shortwave cloud feedback that would be important in the context of understanding climate feedbacks, 1-2 $\mathrm{W\,m^{-2}}$. The ways in which the models are wrong varies, although some patterns emerge. Most models reflect too little solar radiation in the southern storm tracks (cf., Trenberth and Fasullo, 2010a) and misrepresent in one fashion or another the structure of the tropical convergence zones (cf., Lin, 2007). Most models reasonably represent the poleward increase of reflected shortwave radiation in the mid-latitudes of the northern hemisphere, presumably because part of this is carried by the influence of specified surface features such as the Saharan desert and Tibetan Plateau. The models also tend to accurately represent the remarkable constancy in the globally averaged values, and even the year to year variability within latitude bands, but almost all fail to properly capture the near equality of the hemispherically averaged values.

A second example is an evaluation of the simulated trend in global mean temperature over the twentieth century from all of the relevant simulations, 58 in total, in the CMIP3 archive. Fig. 11. Typically temperature changes over the last 100-150 years are presented as an anomaly from a climatological mean. The change in temperature anomaly observed over the twentieth century is accurately reproduced by current climate models (Hegerl et al., 2007). However presentation of the change in temperature anomaly fails to show the differences in representation of global mean temperature, with simulated temperatures at the end of the 20th century exhibiting a range of nearly 3 °C (from 12.8° to 15.5°C). Most models are biased cold, as the multi-model mean temperature is more than 0.5°C colder than measured, an offset that is comparable to the temperature change observed over the 20th Century. From a certain perspective the agreement is excellent; errors in temperature of 1 K out of 288 K corresponds to an error of only 0.35%, albeit somewhat larger (16 $\mathrm{W\,m^{-2}}$ or nearly 7% when translated into an energy flux). However, even such a small temperature error can alter the modeled climate in ways that are as great as the climate change that has occurred over the twentieth century or are projected for the twenty-first century. Such an error has important implications for the surface energy budget and, as well, greatly exceeds the radiative perturbations arising from anthropogenic forcings. The implications of the bias in modeled energy budgets on projected climate change from the considerably smaller radiative forcings is not known. But it is surprising that despite these differences the models, individually and collectively, still represent the trend in twentieth century temperatures as accurately as they do. This surprise is tempered by a realization that the agreement in the 20th century temperature trend may also be a reflection of the model development process and the considerable latitude that uncertainty in the aerosol forcing gives model developers in matching the observed temperature trend (Kiehl, 2007).

The above examples point out departures between models and observations that are typical of current climate models. Some aspects of the climate system are accurately represented, others are not. Those aspects which are well reproduced have often been the target of the model development process, but not always and often indirectly. Importantly, biases in many key energy flows are typically greater than external forcings that have occurred over the past century or that might be expected over the next decades or century and are greater also than changes that might be expected in the response of the climate system to these forcings. The real challenge is to determine the extent to which, despite inaccuracies and artifacts of design decisions, the models can nonetheless be useful. With respect to Fig. 11 it seems useful to ask, given a range of model sensitivities, how great a range of forcings is consistent with understanding of changes in the composition of the atmosphere, and properties of the land surface, over the twentieth century. With regard to Fig. 10 it seems that much insight into the workings of the climate system might be gained by exploring why $\langle Q_n^{\uparrow} \rangle$ varies so little in the models, or what controls the hemispheric differences in planetary albedo, with the idea that the arguments so developed can lead to a more fruitful exploration of observational data. Aspects of the simulated climate that models persistently get wrong, or persistently get right, appear to provide the most profound opportunity to learn new things. Ultimately, however, the merit of any ideas that are developed through such an exercise largely depends on their ability to withstand observational tests. Unfortunately, limitations in our ability to observe the changing energy flows in the climate system severely limit the ability to construct such tests.

5.2 Modeling climate change

We return here to our initial questions: How accurate is the forcing-response paradigm?; What is the equilibrium climate sensitivity?; How and to what extent can compositional perturbations be characterized by a single number, the radiative forcing? We call attention to several new, modeling centered, approaches that have emerged to help advance answers. None of these is perfect, but a few are interesting. For the question of determining S_{eq} a key initiative has been to take advantage of the ready availability of simulation output from many modeling centers in an attempt to articulate and evaluate the physical processes that determine its magnitude. One attractive strategy in this respect is to explore the relationship between model skill and sensitivity across multiple models (using for instance the CMIP archive), or multiple realizations of a single model whose physical parameterizations (usually pertinent to representation of clouds) are perturbed (Murphy et al., 2004). For instance, using the CMIP3 archive, Hall and Qu (2006) showed that the surface albedo feedback, over the seasonal cycle correlates strongly with the surface albedo feedback in a global warming

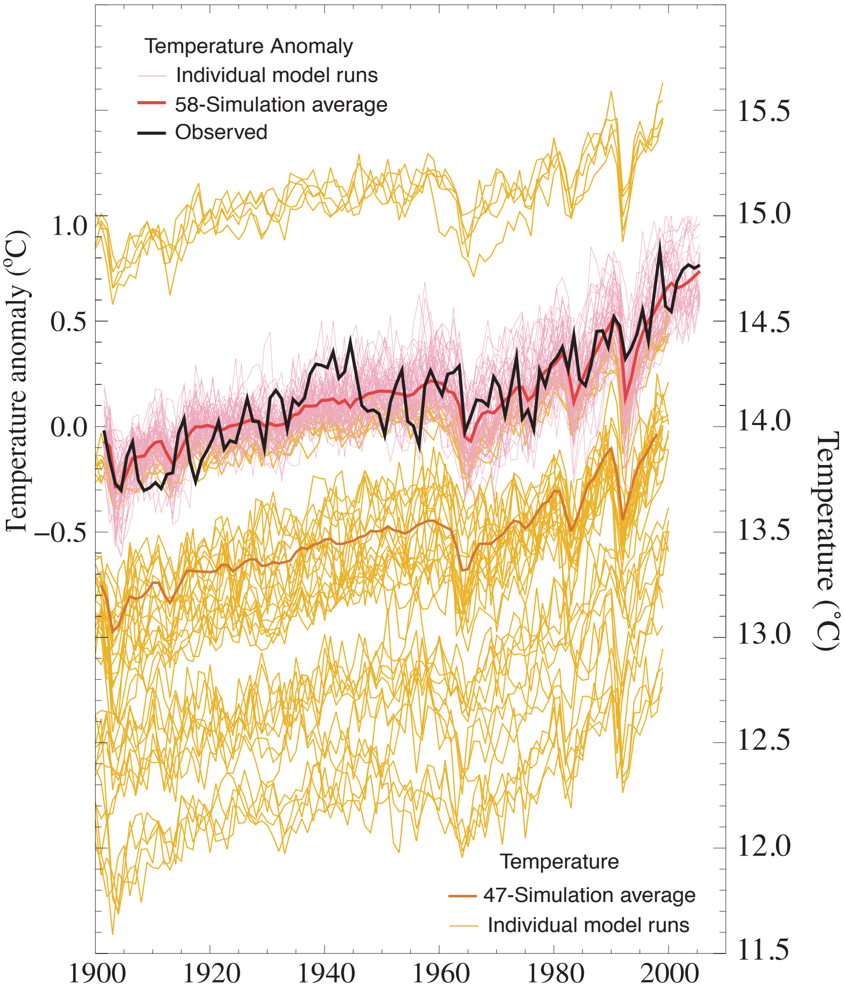

Fig. 11 Change in global mean temperature anomaly (left axis) and global mean temperature (right axis) over the twentieth century as evaluated with climate models that participated in the 2007 IPCC Assessment. The observed change in temperature anomaly (HadCRUT3) is plotted in black. The 14 K offset in the vertical scales corresponds to the 19611990 global mean surface temperature (Jones et al., 1999). Adapted from Tredger (2009) and IPCC AR4 Figure 9.5.

experiment. In so doing they hypothesize that the ability of the model to accurately represent surface albedo feedback on the seasonal cycle is essential if the model is to accurately represent the response of the climate system to external forcings, and that measurements thereof allow inference of the strength of the surface albedo feedback in the climate change context. The key, but usually unstated, assumption is that the relationship between the latter and the former is not an artifact of the model ensemble, but rather a characteristic of the physical system; an assumption which has no *a priori* foundation and often fails in practice Klocke et al. (2011).

Because estimates of aerosol forcing are very dependent on the environment in which the aerosol is found, renewed effort has been devoted to better representing the three-dimensional distribution of size-dependent aerosol composition as a function of time. This information, together with the humidity dependence of the particle size and refractive index can then be used to generate a four-dimensional distribution of aerosol optical properties (extinction coefficient, single scattering albedo, asymmetry parameter) that can be input into radiation transfer models, taking into account solar geometry, surface albedo, and the distribution of clouds to calculate the radiative perturbation of the anthropogenic aerosol (Stier et al., 2011; Bellouin et al., 2008). Although it is unlikely that such complexity is generally necessary, or will ameliorate existing biases in climate models, it would provide a consistent framework for estimating quantities like the direct aerosol radiative effect (DARE) and evaluating how sensitive such estimates are to the underlying climate of the model. Here, however a note of caution is merited. Inter-comparison studies (Kinne et al., 2006) show that although properties such as AOD are rather consistently represented across models, marked differences (factors of three) are seen in how this AOD is attributed to natural versus anthropogenic components, with obvious implications regarding aerosol forcing. Even greater differences are exhibited in diagnostic variables such as the atmospheric residence times of the different chemical substances. This level of disagreement, which may be viewed as a measure of the current state of the art of modeling of tropospheric aerosols, raises questions about the capability at present to model aerosol forcing to the accuracy that is needed to constrain estimates of climate sensitivity.

An important recognition in the past several years is the need to use climate models, rather than radiative transfer calculations, to estimate the forcing from greenhouse gas perturbations. As discussed above, rapid adjustments, particularly those associated with clouds and the stratosphere, can be expected to mediate the effective forcing of greenhouse gases, raising the possibility that estimates of greenhouse gas forcings will also be model dependent. This situation raises the question as to whether the basic climate modeling paradigm that has been employed over the past several decades is the most fruitful approach. To the extent that a procedure could be developed for unambiguously determining the radiative perturbation associated with a change in atmospheric composition, and to the extent which Eq. (1) unambiguously describes the response of the climate system to this forcing, for the purposes of determining S_{eq}, the ultimate source of the forcing (whether it be a change in aerosols, greenhouse gases or surface properties) would not matter.

As evidence mounts that the assumptions underlying the forcing, feedback, response framework are problematic it would seem advantageous if climate model studies adopted an approach where the *radiative* forcing is unambiguous, and the response stands a chance of being testable. In this regard it would seem useful to refocus efforts on an artificial change in the solar irradiance, rather than a change in atmospheric composition. Not only is the radiative forcing known *a priori,* but the eleven year solar cycle, although complicated by its weak amplitude and lack of homogeneity across the solar spectrum, offers an observational analog that could help constrain the modeling. But perhaps in the long run the greatest advantage of such an approach would be that the forcing that is input to the models is well defined and thus that model experiments carried out with different climate models might be more closely compared, allowing inter-model differences to be identified and their causes understood. This approach in turn would help identify physical processes whose model representations can be compared with evolving understanding and thereby lead to improvement in models and understanding.

Concluding remarks

It has long been clear that profound changes in the state of the climate can result from systematic changes in the flows of energy through the system, changes that are small relative to the mean values of these energy flows and even to spatio-temporal variations and fluctuations of these flows on a variety of scales. The implications of this situation on approaches to observing and modeling the Earth system remain unclear. How well can energy flows through the Earth system be measured, or modeled, relative to anticipated changes in the energy balance associated with changing concentrations of greenhouse gases and aerosols? With recognition that models, and observations, are imperfect the question necessarily arises what level of accuracy is required. A further question would address the utility of disaggregations of the system that endeavor to relate changes in energy flows to specific components or processes?

In this article we have reviewed the state of understanding of the energy budget. When due account is taken of present uncertainties there is little evidence of any crisis in understanding of the energy pathways, a finding that is contrary to that of Trenberth and Fasullo (2010b). However the level of uncertainty remains large relative to well known changes resulting from incremental greenhouse gases and to less well known changes resulting from aerosols and to responses of the climate system to these forcings. Measurements likely constrain the top of the atmosphere energy budget to within 4 $\mathrm{W\,m^{-2}}$, but this uncertainty, which is well larger than the total greenhouse gas forcing over the industrial era, about $\mathrm{W\,m^{-2}}$, cannot constrain climate system response to this forcing or to the likely lower total forcing when the forcing by aerosols is folded in. Uncertainty in the surface energy budget, arising mainly from uncertainties in the amount of incident radiation at the surface (both long and shortwave) and precipitation is likely three to five times as large as at the top of the atmosphere.

The principal source of uncertainty in forcing is that associated with the influences of particulate matter, clouds and aerosols, whose high degree of space-time variability and whose strength of coupling to the ambient flow, greatly complicates both the determination of this forcing, and assessment of its ensuant effect on the system. There is a long tradition of trying to isolate the effects of clouds on the energy budget, but this approach seems to introduce as many problems as it solves. Differences between the effects of clouds, versus the effects of cloudy scenes, complicate comparisons between models and observations as well as attempts to attribute changes in the energy budget to clouds, versus other processes. Differences between the fast response of clouds to changes in the atmospheric composition, for instance through perturbations to the ambient aerosol or greenhouse gas concentrations, and the slower response that is due to changes in, surface temperature also complicate efforts to estimate the radiative forcing that can be associated with a perturbation to the system, as distinguished from system feedbacks. These complications reinforce the point that radiative forcing is model dependent and, in large part because of aerosol and cloud effects, are as uncertain as the feedbacks, which determine the net response of the system. However, because both the forcing and the feedbacks are model dependent, attempts to infer basic properties of the climate system, such as its sensitivity to a radiative forcing, from observations alone are likely to be laden with a high degree of uncertainty. Such issues are compounded by the fact that there is no objective definition of a cloud, and considerable differences can emerge as a consequence of differences in measurement approaches, thresholds, and spatial resolution.

In sum, the rich data sets of observations and the many improvements in understanding the processes that comprise Earth's climate system, rather than leading to precise answers to well defined questions are leading rather to a questioning of the fruitfulness of long-standing ways of framing the climate change problem. Does it continue to make sense to differentiate between cloud-free and cloudy scenes? Is the equilibrium climate sensitivity still a useful construct? Can compositional changes be characterized by their radiative forcing? Does the tendency toward unbounded complexity in models accurately reflect improved understanding of the phenomena that important in climate and climate change, or is it possible that phenomena that are just as important are not yet recognized or understood sufficiently well to be represented in climate models? The conventional impulse would be to answer all of these questions in the negative. But this impulse must be tempered by an appreciation that in many cases clearly better alternatives are elusive. While viewing the atmosphere as consisting of discrete, well defined parcels of cloudy and cloud-free air may be comforting and thus far has been central to modeling and observing the climate system, in reality the atmospheric composition (including its particulate component) that controls the distribution of radiation, precipitation, and aerosol processing in three-dimensional space consists much more accurately of an intimate and ephemeral mixture of cloudy and cloud-free air. In this context the very concept of climate sensitivity has thus far remained an unobservable, and even unrealizable construct. Nonetheless determining climate sensitivity poses a theoretical challenge that perhaps should not be dismissed. Is it possible to really understand the climate system or climate change without understanding climate sensitivity? We think not. Ultimately knowledge of Earths climate sensitivity is essential to societal planning of allowable emissions of greenhouse gases and of approaches to meeting the energy needs of human society. For this reason, to separate the difficulty of determining climate sensitivity from the difficulty of relating compositional changes to radiative perturbations we suggest that the approach to determining climate sensitivity might be more usefully framed in terms of perturbations to the solar constant. This approach has the advantage that, given the irradiance fluctuations accompanying the eleven-year sunspot cycle, such changes are, in the long-run, observable. Finally, because the climate sensitivity speaks directly to the question of how particulate matter in the atmosphere mediates externally forced changes, determining climate sensitivity remains the grand challenge of this second century of climate science.

Acknowledgements We thank the organizers of the workshop for giving us the opportunity to present our views. In addition through the development of this manuscript exchanges with Andrea Brose, Robert Cahalan, Robert Cess, Ralph Kahn, Seiji Kato, Norman Loeb, John Lyman, Jochem Marotzke, David Randall, and Graeme Stephens, were a great resource in the development of our ideas and analysis methods. MISR and CERES data were obtained from the NASA Langley Research Center Atmospheric Sciences Data Center. We thank Wei Wu (BNL) for providing cloud cover data for north Central Oklahoma as well as Stefan Kinne and Thorsten Mauritsen for critical and substantive comments on an initial draft version of this manuscript. We acknowledge the modeling groups, the Program for Climate Model Diagnosis and Intercomparison (PCMDI) and the WCRP's Working Group on Coupled Modelling (WGCM) for their roles in making available the WCRP CMIP3 multi-model dataset. Support of this dataset is provided by the Office of Science, U.S. Department of Energy. Work by SES was supported by the US Department of Energy (Office of Science, OBER) under Contract No. DE-AC02-09CH10886. B. Stevens was supported by the Max Planck Society for the Advancement of Science.

References

Abbot, C. G. and F. E. Fowle, 1908: Radiation and terrestrial temperature. *Annals of the Astrophysical Observatory of the Smithsonian institution*, **2 (2)**, 125–224.

Albrecht, B. A., 1989: Aerosols, Cloud Microphysics, and Fractional Cloudiness. *Science*, **245 (4923)**, 1227–1230.

Andrews, T., P. M. Forster, and J. M. Gregory, 2009: A Surface Energy Perspective on Climate Change. *J. Climate*, **22 (10)**, 2557–2570.

Arrhenius, S., 1896: On the influence of carbonic acid in the air upon the temperature of the ground. *Philosophical Magazine*, **41**, 237–276.

Bar-Or, R. Z., I. Koren, and O. Altaratz, 2010: Estimating cloud field coverage using morphological analysis. *Environ. Res. Lett.*, **5 (1)**, 014 022.

Bellouin, N., A. Jones, J. Haywood, and S. A. Christopher, 2008: Updated estimate of aerosol direct radiative forcing from satellite observations and comparison against the Hadley Centre climate model. *J. Geophys. Res.*, **113 (D10)**.

Bender, F. A.-M., H. Rodhe, R. J. Charlson, A. M. L. Ekman, and N. G. Loeb, 2006: 22 views of the global albedo-comparison between 20 GCMs and two satellites. *Tellus A*, **58 (3)**, 320–330.

Box, G. E., 1979: Robustness in the strategy of scientific model building. *Robustness in Statistics: Proceeding of a Workshop*, R. L. Launer and G. N. Wilkinson, Eds., Academic Press, New York, 40.

Brasseur, G. P. and E. Roeckner, 2005: Impact of improved air quality on the future evolution of climate. *Geophys. Res. Lett.*, **32 (23)**.

Caldwell, P. and C. S. Bretherton, 2009: Response of a Subtropical Stratocumulus-Capped Mixed Layer to Climate and Aerosol Changes. *J. Climate*, **22 (1)**, 20–38.

Cess, R. D., 1976: Climate Change: An Appraisal of Atmospheric Feedback Mechanisms Employing Zonal Climatology. *J. Atmos. Sci.*, **33 (10)**, 1831–1843.

Charlson, R. J., S. E. Schwartz, J. M. Hales, R. D. Cess, J. A. Coakley, J. E. Hansen, and D. J. Hofmann, 1992: Climate Forcing by Anthropogenic Aerosols. *Science*, **255 (5043)**, 423–430.

Clement, A. C., R. Burgman, and J. R. Norris, 2009: Observational and Model Evidence for Positive Low-Level Cloud Feedback. *Science*, **325 (5939)**, 460–464.

Clothiaux, E. E., T. P. Ackerman, G. G. Mace, K. P. Moran, R. T. Marchand, M. A. Miller, and B. E. Martner, 2000: Objective Determination of Cloud Heights and Radar Reflectivities Using a Combination of Active Remote Sensors at the ARM CART Sites. *Journal of Applied Meteorology*, **39 (5)**, 645–665.

Colman, R. and B. McAvaney, 2009: Climate feedbacks under a very broad range of forcing. *Geophys. Res. Lett.*, **36 (1)**.

Di Girolamo, L. and R. Davies, 1997: Cloud fraction errors caused by finite resolution measurements. *J. Geophys. Res.*, **102 (D2)**, 1739–1756.

Dines, W. H., 1917: The heat balance of the atmosphere. *Quart. J. Roy. Meteor. Soc*, **43 (12)**, 151–158.

Forster, P., et al., 2007: Changes in Atmospheric Constituents and in Radiative Forcing. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. Averyt, and M. Tignor, Eds., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA., 1–106.

Forster, P. M. F. and J. M. Gregory, 2006: The Climate Sensitivity and Its Components Diagnosed from Earth Radiation Budget Data. *Journal of Climate*, **19 (1)**, 39–52.

Foster, J., S. A. Ackerman, A. K. Heidinger, and B. C. Maddux, 2010: Global cloudiness. Bulletin of the American Meteorological Society, e.

Foster, M., S. A. Ackerman, R. Bennartz, A. K. Heidinger, B. C. Maddux, and W. B. Rossow, 2009: Global Cloudiness [in "State of the Climate in 2008"]. *Bulletin of the American Meteorological Society*, **90 (8)**, S29–S30.

Genkova, I., C. Long, P. Minnis, and P. Heck, 2004: Point-to-Point Comparison of Satellite and Ground-Based Cloud Properties at the ARM Southern Great Plains Central Facility. *Fourteenth ARM Science Team Meeting Proceedings, Albuquerque, New Mexico, March 22-26, 2004*, 1–18.

Gregory, J. and M. Webb, 2008: Tropospheric Adjustment Induces a Cloud Component in CO2 Forcing. *Journal of Climate*, **21 (1)**, 58–71.

Gregory, J. M., R. J. Stouffer, S. C. B. Raper, P. A. Stott, and N. A. Rayner, 2002: An Observationally Based Estimate of the Climate Sensitivity. *Journal of Climate*, **15 (22)**, 3117–3121.

Hall, A. and X. Qu, 2006: Using the current seasonal cycle to constrain snow albedo feedback in future climate change. *Geophys. Res. Lett.*, **33 (3)**, L03 502–4.

Han, Q., W. B. Rossow, J. Zeng, and R. Welch, 2002: Three Different Behaviors of Liquid Water Path of Water Clouds in Aerosol–Cloud Interactions. *J. Atmos. Sci.*, **59**, 726–735.

Hansen, J., M. Sato, and R. Ruedy, 1997: Radiative forcing and climate response. *J. Geophys. Res.*, **102 (D6)**, 6831–6864.

Harries, J. E., et al., 2005: The Geostationary Earth Radiation Budget Project. *Bull. Amer. Meteor. Soc.*, **86 (7)**, 945–960.

Hegerl, G., et al., 2007: Understanding and Attributing Climate Change. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. Averyt, M. Tignor, and H. L. Miller, Eds., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Heidinger, A. K. and M. J. Pavolonis, 2009: Gazing at Cirrus Clouds for 25 Years through a Split Window. Part I: Methodology. *J. Appl. Meteor. Climatol.*, **48 (6)**, 1100–1116.

Holben, B. N., et al., 2001: An emerging ground-based aerosol climatology: Aerosol optical depth from AERONET. *J. Geophys. Res.*, 1–31.

Hunt, G. E., R. Kandel, and A. T. Mecherikunnel, 1986: A history of presatellite investigations of the Earth's Radiation Budget. *Rev. Geophys.*, **24 (2)**, 351–356.

Intrieri, J., M. Shupe, T. Uttal, and B. McCarty, 2002: An annual cycle of Arctic cloud characteristics observed by radar and lidar at SHEBA. *J. Geophys. Res.*, **(C10)**, 1–15.

Jones, P., M. New, D. E. Parker, S. Martin, and I. G. Rigor, 1999: Surface air temperature and its changes over the past 150 years. *Rev. Geophys.*, **37 (2)**, 173–199.

Jung, M., et al., 2010: Recent decline in the global land evapotranspiration trend due to limited moisture supply. *Nature*, **467 (7318)**, 951–954.

Kahn, R. A., B. J. Gaitley, M. J. Garay, D. J. Diner, T. F. Eck, A. Smirnov, and B. N. Holben, 2010: Multiangle Imaging SpectroRadiometer global aerosol product assessment by comparison with the Aerosol Robotic Network. *J. Geophys. Res.*, **115 (D23)**.

Khain, A., 2009: Notes on state-of-the-art investigations of aerosol effects on precipitation: a critical review. *Environ. Res. Lett.*

Kiehl, J. T., 2007: Twentieth century climate model response and climate sensitivity. *Geophys. Res. Lett.*, **34 (22)**.

Kiehl, J. T. and K. E. Trenberth, 1997: Earth's annual global mean energy budget. *Bulletin of the American Meteorological Society*, **78**, 197–208.

Kim, S.-W., S.-C. Yoon, E. Dutton, J. Kim, C. Wehrli, and B. Holben, 2008: Global Surface-Based Sun Photometer Network for Long-Term Observations of Column Aerosol Optical Properties: Intercomparison of Aerosol Optical Depth. *Aerosol Sc. & Tech.*, **42 (1)**, 1–9.

Kinne, S., et al., 2006: An AeroCom initial assessment – optical properties in aerosol component modules of global models. *ACP*, **6 (7)**, 1815–1834.

Klocke, D., R. Pincus, and J. Quaas, 2011: On constraining estimates of climate sensitivity. *J. Climate*, **in press**.

Knutti, R. and G. C. Hegerl, 2008: The equilibrium sensitivity of the Earth's temperature to radiation changes. *Nature Geoscience*, **1 (11)**, 735–743.

Kopp, G. and J. L. Lean, 2011: A new, lower value of total solar irradiance: Evidence and climate significance. *Geophys. Res. Lett.*, **38 (1)**.

Koren, I., G. Feingold, and L. A. Remer, 2010: The invigoration of deep convective clouds over the Atlantic: aerosol effect, meteorology or retrieval artifact? *Atmos. Chem. Phys.*, **10 (18)**, 8855–8872.

Koren, I., L. A. Remer, Y. J. Kaufman, Y. Rudich, and J. V. Martins, 2007: On the twilight zone between clouds and aerosols. *Geophys. Res. Lett.*, **34 (8)**, L08 805.

Levitus, S., J. Antonov, and T. Boyer, 2005: Warming of the world ocean, 1955–2003. *Geophys. Res. Lett.*, **32 (2)**, L02 604–.

Lin, J.-L., 2007: The Double-ITCZ Problem in IPCC AR4 Coupled GCMs: Ocean–Atmosphere Feedback Analysis. *J. Climate*, **20 (18)**, 4497–4525.

Liu, Y. and P. H. Daum, 2002: Anthropogenic aerosols: Indirect warming effect from dispersion forcing. *Nature*, **419 (6907)**, 580–581.

Loeb, N. G., N. Manalo-Smith, S. Kato, W. F. Miller, S. K. Gupta, P. Minnis, and B. A. Wielicki, 2003: Angular Distribution Models for Top-of-Atmosphere Radiative Flux Estimation from the Clouds and the Earth's Radiant Energy System Instrument on the Tropical Rainfall Measuring Mission Satellite. Part I: Methodology. *Journal of Applied Meteorology*, **42 (2)**, 240–265.

Loeb, N. G. and G. L. Schuster, 2008: An observational study of the relationship between cloud, aerosol and meteorology in broken low-level cloud conditions. *J. Geophys. Res.*, **113 (D14)**.

Loeb, N. G., B. A. Wielicki, D. R. Doelling, G. L. Smith, D. F. Keyes, S. Kato, N. Manalo-Smith, and T. Wong, 2009: Toward Optimal Closure of the Earth's Top-of-Atmosphere Radiation Budget. *J. Climate*, **22 (3)**, 748–766.

Loeb, N. G., et al., 2007: Multi-Instrument Comparison of Top-of-Atmosphere Reflected Solar Radiation. *J. Climate*, **20 (3)**, 575–591.

Lohmann, U. and J. Feichter, 2005: Global indirect aerosol effects: a review. *Atmos. Chem. Phys.*, **5**, 715–737.

Lohmann, U., et al., 2010: Total aerosol effect: radiative forcing or radiative flux perturbation? *Atmos. Chem. Phys.*, **10 (7)**, 3235–3246.

London, J. A., 1957: A study of the atmospheric heat balance. Tech. rep.

Long, C. N., T. P. Ackerman, K. L. Gaustad, and J. N. S. Cole, 2006: Estimation of fractional sky cover from broadband shortwave radiometer measurements. *J. Geophys. Res.*, **111 (D11)**, D11 204–.

Manabe, S. and R. T. Wetherald, 1967: Thermal Equilibrium of Atmosphere with a Given Distribution of Relative Humidity. *J. Atmos. Sci.*, **24 (3)**, 241–&.

Markowicz, K. M., P. J. Flatau, A. M. Vogelmann, P. K. Quinn, and E. J. Welton, 2003: Clear-sky infrared aerosol radiative forcing at the surface and the top of the atmosphere. *Q.J.R. Meteorol. Soc.*, **129 (594)**, 2927–2947.

Matthews, H. D. and K. Caldeira, 2007: Transient climate–carbon simulations of planetary geoengineering. *Proceedings of the National Academy of Science*, **104 (24)**, 9949–9954.

Mauritsen, T., et al., 2011: An Arctic CCN-limited cloud-aerosol regime. *Atmos. Chem. Phys.*, **11 (1)**, 165–173.

McComiskey, A., S. E. Schwartz, B. Schmid, H. Guan, E. R. Lewis, P. Ricchiazzi, and J. A. Ogren, 2008: Direct aerosol forcing: Calculation from observables and sensitivities to inputs. *J. Geophys. Res.*, **113 (D9)**, D09 202.

Michalsky, J., F. Denn, C. Flynn, G. Hodges, P. Kiedron, A. Koontz, J. Schlemmer, and S. E. Schwartz, 2010: Climatology of aerosol optical depth in north-central Oklahoma: 1992–2008. *J. Geophys. Res.*, **115 (D7)**.

Mishchenko, M. I., I. V. Geogdzhayev, B. Cairns, W. B. Rossow, and A. A. Lacis, 1999: Aerosol retrievals over the ocean by use of channels 1 and 2 AVHRR data: sensitivity analysis and preliminary results. *Appl. Opt.*, **38 (36)**, 7325–7341.

Murphy, D. M., S. Solomon, R. W. Portmann, K. H. Rosenlof, P. M. Forster, and T. Wong, 2009: An observationally based energy balance for the Earth since 1950. *J Geophys Res-Atmos*, **114**, –.

Murphy, J. M., D. M. H. Sexton, D. N. Barnett, G. S. Jones, M. J. Webb, M. Collins, and D. A. Stainforth, 2004: Quantification of modelling uncertainties in a large ensemble of climate change simulations. *Nature*, **430 (7001)**, 768–772.

Nakajima, T., A. Higurashi, K. Kawamoto, and J. E. Penner, 2001: A possible correlation between satellite-derived cloud and aerosol microphysical parameters. *Geophys. Res. Lett.*, **28 (7)**, 1171–1174.

Nemesure, S., R. Wagener, and S. E. Schwartz, 1995: Direct shortwave forcing of climate by the anthropogenic sulfate aerosol: Sensitivity to particle size, composition, and relative humidity. *J. Geophys. Res.*, **100 (D12)**, 26 105–26 116.

Pierrehumbert, R., 2011: Infrared radiation and planetary temperature. *Physics Today*.

Pincus, R. and M. B. Baker, 1994: Effect of precipitation on the albedo susceptibility of clouds in the marine boundary layer. *Nature*, **372 (6503)**, 250–252.

Plass, G. N., 1956: The Carbon Dioxide Theory of Climatic Change. *Tellus*, **8 (2)**, 140–154.

Quaas, J., O. Boucher, N. Bellouin, and S. Kinne, 2005: Satellite-based estimate of the direct and indirect aerosol climate forcing. *J. Geophys. Res.*, **113**, 9pp.

Quaas, J., et al., 2009: Aerosol indirect effects – general circulation model intercomparison and evaluation with satellite data. *Atmos. Chem. Phys.*, **9 (22)**, 8697–8717.

Ramanathan, V., 1987: The Role of Earth Radiation Budget Studies in Climate and General Circulation Research. *J. Geophys. Res.*, **92 (D4)**, 4075–4095.

Rossow, W. and R. Schiffer, 1999: Advances in understanding clouds from ISCCP. *Bulletin of the American Meteorological Society*, **80 (11)**, 2261–2287.

Rossow, W. B. and B. Cairns, 1995: Monitoring changes of clouds. *Climatic Change*, **31 (2)**, 305–347.

Schneider, S. H., 1972: Cloudiness as a Global Climatic Feedback Mechanism: The Effects on the Radiation Balance and Surface Temperature of Variations in Cloudiness. *J. Atmos. Sci.*, **29 (8)**, 1413–1422.

Schwartz, S. E., 2011: Feedback and sensitivity in an electrical circuit: an analog for climate models. *Climatic Change*, **106 (2)**, 315–326.

Schwartz, S. E., R. J. Charlson, R. A. Kahn, J. A. Ogren, and H. Rodhe, 2010: Why Hasn't Earth Warmed as Much as Expected? *J. Climate*, **23 (10)**, 2453–2464.

Schwartz, S. E., Harshvardhan, and C. M. Benkovitz, 2002: Influence of anthropogenic aerosol on cloud optical depth and albedo shown by satellite measurements and chemical transport modeling. *Proceedings of the National Academy of Sciences*, **99 (4)**, 1784–1789.

Segrin, M. S., J. A. J. Coakley, and W. R. Tahnk, 2007: MODIS observations of ship tracks in summertime stratus off the west coast of the United States. *J. Atmos. Sci.*, **64 (12)**, 4330–4345.

Shindell, D. and G. Faluvegi, 2009: Climate response to regional radiative forcing during the twentieth century. *Nature Geoscience*, **2 (4)**, 294–300.

Smirnov, A., et al., 2009: Maritime Aerosol Network as a component of Aerosol Robotic Network. *J. Geophys. Res.*, **114 (D6)**, D06 204.

Soden, B. J., I. M. Held, R. Colman, K. M. Shell, J. T. Kiehl, and C. A. Shields, 2008: Quantifying Climate Feedbacks Using Radiative Kernels. *J. Climate*, **21 (14)**, 3504–3520.

Sohn, B. J., T. Nakajima, M. Satoh, and H. S. Jang, 2010: Impact of different definitions of clear-sky flux on the determination of longwave cloud radiative forcing: NICAM simulation results. *Atmos. Chem. Phys.*, **10 (23)**, 11 641–11 646.

Stackhouse, P. W., S. K. Gupta, S. J. Cox, T. Zhang, J. C. Mikovitz, and L. M. Hinkelman, 2011: 24.5 year surface radiation budget data set relaseased. *GEWEX News*, **21 (1)**, 1–4.

Stephens, G. L., et al., 2008: CloudSat mission: Performance and early science after the first year of operation. *J. Geophys. Res.*, **113**.

Stevens, B. and G. Feingold, 2009: Untangling aerosol effects on clouds and precipitation in a buffered system. *Nature*, **461 (7264)**, 607–613.

Stevens, B., et al., 2003: On entrainment rates in nocturnal marine stratocumulus. *Q.J.R. Meteorol. Soc.*, **129 (595)**, 3469–3493.

Stier, P., et al., 2011: The aerosol climate model ECHAM5-HAM. *Atmos. Chem. Phys.*, **5**, 1125–1156.

Tackett, J. L. and L. Di Girolamo, 2009: Enhanced aerosol backscatter adjacent to tropical trade wind clouds revealed by satellite-based lidar. *Geophys. Res. Lett.*, **36 (14)**, L14 804–.

Tredger, E., 2009: On the evaluation of uncertainty in climate models. Ph.D. thesis, London School of Economics, London.

Trenberth, K. E. and J. T. Fasullo, 2010a: Simulation of Present-Day and Twenty-First-Century Energy Budgets of the Southern Oceans. *J. Climate*, **23 (2)**, 440–454.

Trenberth, K. E. and J. T. Fasullo, 2010b: Tracking Earth's Energy. *Science*, **328 (5976)**, 316–317.

Trenberth, K. E., J. T. Fasullo, and J. Kiehl, 2009: Earth's Global Energy Budget. *Bull. Amer. Meteor. Soc.*, **90 (3)**, 311–323.

Twohy, C. H., J. A. Coakley Jr., and W. R. Tahnk, 2009: Effect of changes in relative humidity on aerosol scattering near clouds. *J. Geophys. Res.*, **114 (D5)**, D05 205.

Twomey, S., 1974: Pollution and planetary albedo. *Atmospheric Environment*, **8**, 1251–1256.

Weatherhead, E. C., et al., 2000: Detecting the recovery of total column ozone. *J. Geophys. Res.*, **105 (D17)**, 22 201–22 210.
Wielicki, B. A., B. R. Barkstrom, E. F. Harrison, R. B. Lee, G. Louis Smith, and J. E. Cooper, 1996: Clouds and the Earth's Radiant Energy System (CERES): An Earth Observing System Experiment. *Bulletin of the American Meteorological Society*, **77 (5)**, 853–868.
Wielicki, B. A. and R. M. Welch, 1986: Cumulus Cloud Properties Derived Using Landsat Satellite Data. *Journal of Climate and Applied Meteorology*, **25 (3)**, 261–276.
Winker, D. M., W. H. Hunt, and M. J. McGill, 2007: Initial performance assessment of CALIOP. *Geophys. Res. Lett.*, **34 (19)**, L19 803–.
Zhang, J. and J. S. Reid, 2010: A decadal regional and global trend analysis of the aerosol optical depth using a data-assimilation grade over-water MODIS and Level 2 MISR aerosol products. *Atmos. Chem. Phys.*, **10 (22)**, 10 949–10 963.